SPATIOTEMPORAL MODELING OF BRAIN DYNAMICS USING MACHINE LEARNING APPROACHES

A Dissertation
Presented to
The Academic Faculty

by

Shiyang Chen

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Biomedical Engineering

Georgia Institute of Technology and Emory University
December 2017

SPATIOTEMPORAL MODELING OF BRAIN DYNAMICS USING MACHINE LEARNING APPROACHES

Approved by:

Dr. Xiaoping Hu, advisor
Department of Bioengineering
University of California, Riverside

Dr. Shella Keilholz, advisor
Department of Biomedical Engineering
Georgia Institute of Technology and Emory University

Dr. Chin-Hui Lee
Department of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Krish Sathian
Department of Neurology
Pennsylvania State University

Dr. Lena Ting
Department of Biomedical Engineering
Georgia Institute of Technology and Emory University

Dr. Deqiang Qiu
Department of Biomedical Engineering
Georgia Institute of Technology and Emory University

Date Approved: July 24, 2017

ACKNOWLEDGEMENTS

First of all, I would like to especially thank my advisors, Dr. Xiaoping Hu, and Dr. Shella Keilholz for the patience and support. I would also like to express my gratitude to my committee, Dr. Chin-Hui Lee, Dr. Krish Sathian, Dr. Lena Ting, and Dr. Deqiang Qiu for their comments and suggestions on my research, and I would like to express my gratitude to Dr. Bruce Crosson for his support during the past year.

Next, I would like to thank all my lab members for the past few years. I am fortunate to work in biomedical imaging technology center (BITC), magnetic resonance imaging of neural dynamics (MIND) laboratory, and computational neuroimaging and neuroscience laboratory (CN2L) and I have learned a lot from my colleagues.

Finally, I would like to thank my family, especially my fiancée, Hao Wu, and my parents, Tongxin Chen and Yueqiu Tian.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

LIST OF TABLES vi

LIST OF FIGURES vii

LIST OF SYMBOLS AND ABBREVIATIONS xi

SUMMARY xiii

CHAPTER 1. Introduction 1
1.1 Significance of the study 1
1.2 Thesis organization 3

CHAPTER 2. Spatiotemporal modeling of brain dynamics using resting-state fMRI with Gaussian hidden Markov model 5
2.1 Introduction 5
2.2 Materials and methods 8
2.2.1 Dataset and preprocessing 8
2.2.2 Brain state switching model 8
2.2.3 Stability and reproducibility 11
2.2.4 Brain state sequence decoding 13
2.2.5 Spontaneous brain activity pattern of each state 13
2.3 Results 13
2.4 Discussion 22
2.5 Conclusion 27

CHAPTER 3. Reproducibility of brain States Detected by GHMM of and its Applications 29
3.1 Introduction 29
3.2 Materials and Methods 32
3.2.1 Dataset 32
3.2.2 Preprocessing 33
3.2.3 Training GHMM 34
3.3 Results 35
3.3.1 Reproducibility of brain states on different datasets 35
3.3.2 Brain states under different preprocessing steps 37
3.3.3 PAE 44
3.4 Discussion 51
3.4.1 Reproducibility on different datasets 52
3.4.2 Effect of temporal filtering 53
3.4.3 Effect of GSR 53
3.4.4 Impact of different atlases 54
3.4.5 Effect of PAE 55

3.5 Conclusion **57**

CHAPTER 4. Individual identification using functional brain fingerprint detected by recurrent neural network **58**
4.1 Introduction **58**
***4.2* Materials and methods** **59**
4.2.1 Dataset and preprocessing 59
4.2.2 Recurrent learning model 60
4.2.3 Impact of preprocessing 62
4.3 Results **63**
4.4 Discussion **69**
4.4.1 Duplications of GRU patterns 70
4.4.2 Effect of anatomical differences 70
4.4.3 Effect of physiological signal 72
4.4.4 Limitations and future directions 73
4.5 Conclusion **74**

CHAPTER 5. Conclusion and future work **75**

LIST OF TABLES

Table 1 – The validation accuracy and test accuracy of three different processing approaches. 63

Table 2 – The number of GRUs that capture +DMN-TPN and +TPN-DMN patterns under 3 different preprocessing approaches. 68

Table 3 – The test accuracy on resting state and language task fMRI of three different processing approaches. 72

LIST OF FIGURES

Figure 1 – A two state toy example of the brain's GHMM. In this case, the brain is only switching between two states, introspection and external processing. The numbers on the arrows represent the probabilities of state switching. During the introspection state, the DMN is activated, while in the external processing state, activation is assumed to be in the dorsal attention network. 9

Figure 2 – The schematic of state matching and sorting (A) and the resulting stability of each brain state when the total number of states is set to 9 (B), 10 (C). A toy example of the schematic of the state matching and sorting method with three states and three realizations are shown in (A). Spatial patterns of the brain states and the matched groups are represented by shapes and colors, respectively. The stability of brain states for M=9 and M=10 is shown in (B) and (C). Note that when M=9, brain states from all the realizations are matched perfectly into 9 groups, while when M=10, they are not matched perfectly but grouped into 11 groups instead. 14

Figure 3 – The z-maps of 9 reproducible brain states. The color bar is set to only display z-scores below -20 or above 20. These brain states are different combinations of activated and deactivated brain regions. 15

Figure 4 – Temporal characteristics of the brain state. (A) The state sequence and state probability during one subject scan. Only a period of the scan is demonstrated in the figure for better visualization. Different colors indicate what the subject's brain state is at each time point. The black curves are the posterior probability of different brain states. (B) Boxplots of states' occurrence and average duration. Note that the length of each scan is 14.4 minutes. 17

Figure 5 – The transition properties of the brain states. (A) Average direct transition per scan. Note that only when the transition times from state i to state j is significantly higher than average ($p<0.0007$), the block in the *ith* row and *jth* column will be color coded. (B) Histogram of the duration of the motif from the beginning of S7 to the end of S6. 18

Figure 6 – The reproducibility of brain states for split-half samples. (A) The z-map of brain states detected from all 100 subjects and two non-overlapping halves of the dataset. States with similar spatial patterns are aligned in the same column. (B) The spatial correlation coefficients of two states from the two halves are plotted to show 19

the similarity between them. Note that the last state has the lowest spatial similarity and the lowest z-score on its z-map.

Figure 7 – The reproducibility of the temporal characteristics of brain states for split-half samples. (A) The boxplot of time occupation from two halves; (B) the boxplot of occurrence from two halves; (C) the boxplot of average duration from two halves. The star marks the significant difference between two halves. 20

Figure 8 – The reproducibility of cluster centers derived by k-means. (A) The z-map of cluster centers detected from all 100 subjects and two non-overlapping halves of the dataset. Similar spatial patterns are aligned in the same column. (B) The spatial correlation coefficients of two cluster centers from the two halves are plotted to show the similarity between them. (C) Spatial correlation coefficients between states derived by GHMM and cluster centers derived by k-means. 21

Figure 9 – Whole brain axial image for S6, S7, and S9. 26

Figure 10 – The reproducibility of brain states' spatial patterns. Z-maps of (A) the brain states derived based on the HCP dataset and (B) the brain states derived based on the Beijing_Zang dataset. 36

Figure 11 – The reproducibility of the temporal characteristics of brain states between two datasets (Left: HCP; right: Beijing_Zang). (A) The boxplot of time occupation; (B) the boxplot of occurrence; (C) the boxplot of average duration. The star marks the significant difference between two datasets. 37

Figure 12 – The spatial patterns of brain states before (A) and after (B) temporal filtering (0.01-0.1Hz). Note that no GSR was performed and 236 ROIs were used to extract each time series. The state number here does not imply they have one-to-one correspondence before and after temporal filtering. It is only for visualization and comparison **purposes.** 39

Figure 13 – The temporal characteristics between the corresponding brain states derived before (left) and after (right) temporal filtering. The temporal characteristics are compared using boxplots of (A) time occupation, (B) occurrence, and (C) average duration. 40

Figure 14 – The spatial patterns of brain states under three different preprocessing approaches. (A) Data extracted from atlas with 236 ROIs, without GSR; (B) atlas with 236 ROIs, with GSR; (C) atlas with 360 ROIs, with GSR. 42

Figure 15 – The spatial similarity between the corresponding brain states derived based on 236 ROIs and 360 ROIs. 43

Figure 16 – The temporal characteristics between the corresponding brain states derived based on 236 ROIs (left) and 360 ROIs (right). The temporal characteristics are compared using boxplots of (A) time occupation, (B) occurrence, and (C) average duration. The star marks the significant difference between results from two atlases. 44

Figure 17 – The spatial patterns of brain states for control and PAE group. Note that only those voxels that are significantly activated or deactivated are color coded in the figure (p<0.05/total number of voxels). 46

Figure 18 – The temporal characteristics between the corresponding brain states derived from control (left) and PAE (right) groups. The temporal characteristics are compared using boxplots of (A) time occupation, (B) occurrence, and (C) average duration. 47

Figure 19 – The spatial patterns of brain states for the control and PAE groups. The difference between the groups is shown in the 3rd column. Note that only those voxels that are significantly activated, deactivated, or differences between groups are color coded in the figure (p<0.05/total number of voxels**)**. 48

Figure 20 – The temporal characteristics between the corresponding brain states derived from control (left) and PAE (right) groups. The temporal characteristics are compared using boxplots of (A) time occupation, (B) occurrence, and (C) average duration. The stars indicate a significant difference between brain states (p<0.05/21). 49

Figure 21 – The spatial patterns of brain states for control and PAE groups. The difference between the groups is shown in the 3rd column. Note that only those voxels that are significantly activated, deactivated, or differences between groups are color coded in the figure. The Benjamini-Hochberg procedure was adopted for multiple comparisons (α = 0.05**).** 50

Figure 22 – The spatial patterns of brain states for control and PAE groups with 9 brain states. The difference between the groups is shown in the 3rd column. Note that only those voxels that are significantly activated, deactivated, or differences between groups are color coded in the figure. The Benjamini-Hochberg procedure was adopted for multiple comparisons (α = 0.05**).** 51

Figure 23 –Two atlases with 236 and 360 ROIs. 55

Figure 24 – The architecture of our three-layer recurrent neural network. A 25% dropout rate is applied to the input of average pooling layer as well as the input and recurrent connections of GRU layer. 62

Figure 25 – Test data from different subjects are clustered as the output of the GRU. Dots with different colors are from different subjects. The 256-dimensional output of GRU is visualized in 2 dimensions by t-SNE. 64

Figure 26 – The output of one GRU over time. Each row of the image represents the output of GRU with respect to a different subject. Each column of the image is the output of GRU at a particular time point. 65

Figure 27 – Five representative GRU patterns that resemble 5 different RSNs 65

Figure 28 – Dendrogram of hierarchical clustering to group 256 GRU patterns into 5 clusters. The red line indicates the threshold to separate all GRU patterns into clusters. 66

Figure 29 – Cluster centers of GRU patterns with 3 different preprocessing approaches. The number below the image indicates the number of GRU patterns assigned to each cluster. 67

Figure 30 – Cluster centers of GRU patterns with 3 different preprocessing approaches. The number below the image indicates the number of GRU patterns assigned to each cluster. Only clusters with more than 10 GRU patterns are shown here. 68

Figure 31 – Cluster centers of GRU patterns for gender prediction. The number below the image indicates the number of GRU patterns assigned to each cluster. Only clusters with more than 10 GRU patterns are shown here. 69

LIST OF SYMBOLS AND ABBREVIATIONS

CAPs	co-activation patterns
ICA	independent component analysis
FC	functional connectivity
RNN	recurrent neural network
GRU	gated recurrent unit
GHMM	Gaussian hidden Markov model
fMRI	functional magnetic resonance imaging
FWHM	resting state fMRI
HCP	human connectome project
GSR	global signal regression
ROI	region of interest
TPN	task positive network
FPN	fronto-parietal network
DMN	default mode network
DAN	dorsal attention network
CON	cingulo-opercular network
SMN	somatosensory motor network
VN	visual network
CPAC	configurable pipeline for the analysis of connectomes
HMM	hidden Markov model
TFM	temporal functional mode
FWHM	full width at half maximum

PCA	principle component analysis
PAE	prenatal alcohol exposure
FAS	fetal alcohol syndrome

SUMMARY

Resting state fMRI (rfMRI) has been widely used to study functional connectivity of human brains. Although most of the analysis methods to date have assumed the resting state to be stationary, this assumption is invalid when significant changes in functional connectivity occur within a short period of time, as indicated by recent literature. Hence, resting state analysis can be improved by accounting for temporal changes. Recent literature has reached a consensus that the brain is frequently switching among a number of quasi-stable states. To capture these temporal dynamics and derive both spatial and temporal characteristics of the states simultaneously, we propose two machine learning models that incorporate spatial and temporal information. First, Gaussian hidden Markov model (GHMM) has been applied to systematically model the brain state switching processing in rfMRI. We have shown its stability and reproducibility on healthy controls and then demonstrate its application in deriving biomarkers of prenatal alcohol exposure. Second, we have introduced recurrent neural network to investigating individual uniqueness using fMRI and visualized the features that can distinguish individuals, which is considered as a fingerprint of the brain's function.

CHAPTER 1. INTRODUCTION

1.1 Significance of the study

Since Biswal and colleagues (1995) discovered spontaneous activities in resting-state functional magnetic imaging (rfMRI), rfMRI has been widely used to investigate functional connectivity of the brain without requiring subjects to perform tasks. Most of the rfMRI analysis approaches to date assume that the brain is stationary during the entire scan and treat each time acquisition as different samples of the brain, e.g. full-scan functional connectivity analysis. Some of these approaches further rely on assumptions of independence, orthogonality or sparsity to derive different components of the brain, such as spatial independent component analysis (ICA) (Beckmann, DeLuca et al. 2005), principle component analysis (PCA) and dictionary learning (Lv, Jiang et al. 2013). However, the stationary assumption of resting state was shown to be invalid by Allen and colleagues (2014) demonstrating significant changes in functional connectivity patterns within a short period of time.

Instead of assuming the brain to be stationary, a consensus has been reached that the brain is constantly switching among a number of brain states even when a subject is at rest. Sliding window based approaches have been proposed to capture changes in functional connectivity due to switching brain states in rfMRI (Allen, Damaraju et al. 2014). The fixed window length in these approaches (Hutchison, Womelsdorf et al. 2013, Shakil, Lee et al. 2016) typically ranges from tens of seconds to a few minutes. However, a predefined fixed window is controversial because it may lead to contamination of signal from multiple states and even cancellation of signals. In addition, some approaches

ignore the temporal order of rfMRI data, e.g. co-activation patterns (CAPs), which means that temporally shuffling the data will not change the results they derived (Liu, Chang et al. 2013, Liu and Duyn 2013). We, in contrast, believe that brain states can be better understood by studying the sequential information in fMRI data.

Repeated temporal brain activity patterns have been reported in the literature at different temporal scales. For example, the lag thread method (Mitra, Snyder et al. 2015) decomposes fMRI data into 8 different spatiotemporal patterns with a duration of 2 seconds. Another example is the quasi-periodic pattern, which is a 20-second video of brain activity that frequently occurs in both human and rat brains at rest (Majeed, Magnuson et al. 2011, Thompson, Pan et al. 2014). In order to systematically model and fully utilize the temporal information in fMRI data, we have proposed to apply two machine learning models that can capture both the spatial and temporal characteristics of the brain. These are the Gaussian hidden Markov model (GHMM), and recurrent neural network (RNN).

We have introduced GHMM to model state switching processes of the brain. This model has two parts; the underlying Markov process, and the multivariate Gaussian distribution. The Markov process models the process of a brain switching from one state to another. Given the current brain state, which state is more likely to be the next is trained based on the temporal information from fMRI data. The multivariate Gaussian distribution on the other hand, models the spatial activation pattern of the brain in each state. By utilizing both spatial and temporal information from the fMRI data, we believe that GHMM is a more systematic way of modeling the brain. GHMM can also be applied to detect functional biomarkers of brain diseases, e.g. prenatal alcohol exposure (PAE),

allowing for spatial and temporal features of the brain states to be extracted and compared between groups.

RNN has been used to study the individual uniqueness of brain function. Researchers are aware of the differences in each individual's brain structure, and several brain structural templates has been developed to normalize each brain so that group level analysis can be conducted (Talairach and Tournoux 1988, Mazziotta, Toga et al. 1995). However, after registering each subject's brain to a template, the time courses from the same brain region are often concatenated and averaged out in group analysis. A recent study shows that each subject has a unique brain functional connectivity pattern which can be used to predict the subject's identity (Finn, Shen et al. 2015). However, they have also shown that this functional connectivity pattern is less robust when only given data from a short period of time. We hypothesize that the lack of robustness is because functional connectivity does not fully utilize the temporal information in fMRI data. When computing functional connectivity, the temporal axis collapses and only a spatial pattern is left. Therefore, we propose to use RNN which takes fMRI data in recurrently so that all the details in temporal axis can be considered by the model.

1.2 Thesis organization

In chapter 2, to systematically model spatiotemporal patterns of the brain, GHMM has been adopted to model the brain state switching process. We assumed that the brain switches among a number of different brain states as a Markov process and used multivariate Gaussian distributions to represent the spontaneous activity patterns of brain states. This model has been applied to rfMRI data from 100 subjects in the human

connectome project and detected 9 stable and highly reproducible brain states, including spatial patterns as well as temporal and transition characteristics. Our results indicate that the GHMM can unveil brain dynamics that may provide additional insights regarding the brain at resting state.

In Chapter 3, we have further tested the reproducibility of GHMM on different datasets and studied how different preprocessing steps such as temporal filtering, global signal regression (GSR), and extracting time courses from different atlases, could affect the brain states detected by GHMM. We have found that GHMM is highly reproducible, even on different datasets and with different preprocessing approaches. We applied the model to investigate functional biomarkers for PAE, and have shown that PAE affects the brain switching process on a system level and weakens the anti-correlation between the default mode network and task positive network.

In Chapter 4, we introduce a recurrent neural network based model that can identify individuals with only a short segment of data. In addition, we have demonstrated how global signal and different atlases affected the individual discriminating power, and which neuronal features are important in terms of individual identification. Our results indicate that features important for individual uniqueness provide additional insights regarding human brain dynamics.

CHAPTER 2. SPATIOTEMPORAL MODELING OF BRAIN DYNAMICS USING RESTING-STATE FMRI WITH GAUSSIAN HIDDEN MARKOV MODEL

2.1 Introduction

Understanding the human brain as a dynamical system is gaining traction in the literature (Fox, Snyder et al. 2005, Rabinovich and Muezzinoglu 2010), where the human brain is assumed to frequently switch among different metastable states instead of lingering in a single state. Moreover, dynamic approaches, such as sliding window correlation, have shown advantages over stationary ones in detecting neurological diseases, such as schizophrenia (Sakoğlu, Pearlson et al. 2010). These advantages arise from additional features gained by analyzing the brain signals dynamically. Therefore, modeling brain state switching as a dynamic system can unveil additional characteristics about the underlying processes of human brains.

Existing dynamic analysis approaches have limitations. For example, some techniques, such as co-activation patterns (CAPs) (Liu, Chang et al. 2013, Liu and Duyn 2013), spatial independent component analysis (ICA) (Beckmann, DeLuca et al. 2005), and temporal ICA (Smith, Miller et al. 2012) cannot fully exploit information contained in the temporal order of functional magnetic resonance imaging (fMRI) time frames. When the temporal order of fMRI time frames is ignored, each time frame is treated as an independent sample of the brain, and thus shuffling fMRI time frames does not affect the spatial patterns derived by these approaches. In addition, ICA-based approaches assume

independence of components and rely on this assumption to derive spatial patterns. Other dynamic methods take the sequential order of fMRI time frames into consideration. For example, the sliding window approach, which has been applied to estimate fluctuations in functional connectivity (Allen, Damaraju et al. 2014), retains sequential information in the data. However, because the length of the sliding window is fixed (Keilholz 2014), signals from multiple states may be mixed in each window, resulting in contamination between states and maybe even cancellation of signals. However, while these methods present simple statistics, such as the occurrence rate of CAPs or functional connectivity states, quantification of the sequential transitions between CAPs or states is not embedded in these models.

In contrast, we focus on quantifying the sequential transitions between different brain states using transition probabilities and model the brain with a Gaussian hidden Markov model (GHMM) (Eddy 1996, Bilmes 1998). Assuming the brain transits among different states over time, our GHMM models the state switching process as a Markov chain and the spontaneous activity pattern of each brain state as a multivariate Gaussian distribution. Unlike ICA, which assumes the components to be independent of each other, our model makes no hypothesis on the relationship between brain states.

The hidden Markov model (HMM) has been previously employed as a sequential modeling tool in studying the brain (Jones, Fontanini et al. 2007, Eavani, Satterthwaite et al. 2013, Baker, Brookes et al. 2014, Ou, Xie et al. 2014). The HMM was applied to electrophysiological data and detected 4 brain states of neuronal firing patterns in rodents subjected to different types of stimuli (Jones, Fontanini et al. 2007). Applied to human data, the HMM revealed temporal variability of functional connectivity in

magnetoencephalography (Baker, Brookes et al. 2014) and fMRI (Eavani, Satterthwaite et al. 2013, Ou, Xie et al. 2014). A generalized HMM, the hidden Markov random field, was applied in fMRI to detect binary state (on/ off) switching on a voxel level (Lindquist, Waugh et al. 2007, Robinson, Wager et al. 2010, Liu, Awate et al. 2014).

Compared to the previous applications of HMM in fMRI, our approach is substantially different. The model in (Ou, Xie et al. 2014) was based on the functional connectivity strength derived by sliding window method while our model is directly based on fMRI time courses, which will not encounter the aforementioned blurring by the use of sliding window approach. Although Eavani, et al (Eavani, Satterthwaite et al. 2013) also modeled fMRI time courses using an HMM, they studied functional connectivity changes using covariance matrices from each state. In contrast, we investigate temporal dynamics of spontaneous brain activity by analyzing the average activity pattern of each state. The hidden Markov random field in (Lindquist, Waugh et al. 2007, Robinson, Wager et al. 2010, Liu, Awate et al. 2014) assumed that each voxel had only two states (on/ off), while our GHMM does not have an assumption on the number of states. Furthermore, the criteria for determining the number of states in previous work was relatively ambiguous with the number of states determined by the elbow point (Eavani, Satterthwaite et al. 2013) or local optimum(Ou, Xie et al. 2014) of the model's log-likelihood. Here, we introduce a more quantitative and robust method of using stability as a criterion for determining the number of states.

To the best of our knowledge, GHMM has not hitherto been applied to model the sequential state switching process of spontaneous brain activities. The GHMM quantifies the sequential transition of brain states and is not restricted by the aforementioned

limitations of CAPs, ICA, and sliding window. In this paper, we describe this approach and its application in a comprehensive analysis of brain spontaneous activities based on publicly available human connectome project (HCP) dataset.

2.2 Materials and methods

2.2.1 Dataset and preprocessing

One hundred subjects from HCP (Van Essen, Smith et al. 2013) S500 release were used in the present work (age: 22-36^{+}, gender: 46M/54F, TR=0.72 s, scan duration: 14.4 minutes, one scan for each subject). Preprocessing was performed according to the minimal preprocessing pipelines of HCP (Glasser, Sotiropoulos et al. 2013). In particular, we considered the data that had been processed with ICA-FIX (Griffanti, Salimi-Khorshidi et al. 2014, Salimi-Khorshidi, Douaud et al. 2014), registered onto 32k Cento69 surface mesh (Van Essen, Glasser et al. 2012) and slightly smoothed with 2 mm full width at half maximum (FWHM) kernel. All time courses were further filtered with a 0.01–0.1 Hz band-pass filter.

In order to reduce computational complexity, signals from 236 regions of interest (ROIs) (Power, Cohen et al. 2011) were extracted by averaging time courses of voxels within a 5 mm radius. This distance was calculated based on surface distance by Dijkstra's algorithm (Hanke, Halchenko et al. 2009). The mean value of each time course was subtracted and the standard deviations were normalized to 1 before these time courses were fed into the GHMM.

2.2.2 Brain state switching model

To model resting-state fMRI time courses in a non-stationary manner, we applied a sequential modeling tool, GHMM, to analyze the data. It has been shown that the brain is constantly switching from one metastable state to another (Rabinovich and Muezzinoglu 2010). In this work, the sequential switching process of brain states was modeled as a Markov chain, while an observation of a brain state (i.e. an fMRI time acquisition) was modeled by a multivariate Gaussian distribution. A toy example illustrating the use of a GHMM with two states is shown in Figure 1. In this example, the brain is switching between two different brain states (introspection state and external processing state), and the observation from fMRI are activations in default mode network (DMN) and dorsal attention network, respectively.

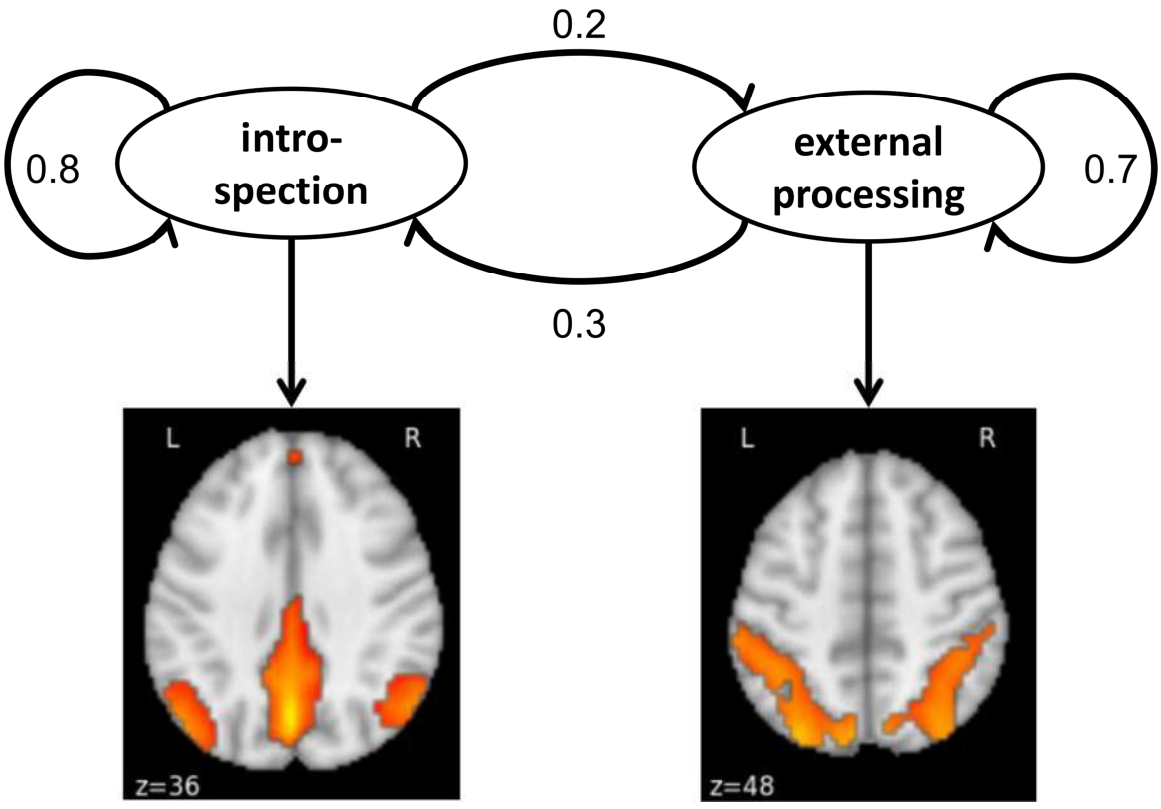

Figure 1 – A two state toy example of the brain's GHMM. In this case, the brain is only switching between two states, introspection and external processing. The numbers

on the arrows represent the probabilities of state switching. During the introspection state, the DMN is activated, while in the external processing state, activation is assumed to be in the dorsal attention network.

In order to derive the GHMM that best fits the fMRI observations, its parameter set $\tilde{\lambda}$ was estimated using the following objective function:

$$\tilde{\lambda} = \arg\max_{\lambda} L(\lambda; \boldsymbol{O}), \tag{1}$$

where λ is the parameter set for the GHMM; $\boldsymbol{O} = \{\boldsymbol{O}_t\}, t = 1, 2, \dots, T$, are the observations from fMRI and T denotes the total number of fMRI acquisitions per person. $L(\lambda; \boldsymbol{O})$ is the likelihood of obtaining the fMRI observations $\boldsymbol{O}$ given that the parameter set is λ. In the GHMM, the probability of the observations under a state is modeled as a multivariate Gaussian distribution. In our case, the mean vector of the Gaussian distribution represents the average activation pattern of each brain state, and its covariance matrix is used to model the variability within each brain state. With a Gaussian assumption, the objective function can be written as:

$$L(\lambda; \boldsymbol{O}) = \sum_{q_0, q_1, \dots, q_T} P(q_0) \prod_{t=1}^{T} a_{q_{t-1} q_t} \frac{1}{\sqrt{(2\pi)^N \left|\boldsymbol{\Sigma}_{q_t}\right|}} exp\left(-\frac{1}{2}(\boldsymbol{O}_t - \boldsymbol{\mu}_{q_t})^T \boldsymbol{\Sigma}_{q_t}^{-1} (\boldsymbol{O}_t - \boldsymbol{\mu}_{q_t})\right), \tag{2}$$

the parameters in the above equation are as follows. $a_{ij} = P(q_{t-1} = i | q_t = j)$ denotes the probability of switching from state i to state j. $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ are the mean vector and covariance matrix of the multivariate Gaussian distribution under state i. $\lambda =$

$\{a_{ij}, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}, i, j = 1, 2, \ldots, M$, where M is the total number of hidden brain states. N ($N = 236$) denotes the total number of time courses that are fed into the model and $P(q_0)$ represents the initial state probability. $P(q_1) = \sum_{q_0} P(q_0) a_{q_0 q_1}$ is assumed to be a uniform distribution such that q_1 can start with any states with equal probability. To further reduce computational complexity, we set the covariance matrix of the multivariate Gaussian to be diagonal, i.e., each dimension of the Gaussian distribution is uncorrelated, assuming that the signal changes (noise and individual variability) in different ROIs at each state are uncorrelated.

The observations from fMRI time courses, $\boldsymbol{O}$, were fed into a k-means clustering program to identify M cluster centers, and the resultant cluster centers were used to initialize the mean of the Gaussian distribution, $\boldsymbol{\mu}_i$. The covariance matrices of the Gaussian, $\boldsymbol{\Sigma}_i$, were initialized as the covariance matrix of the first subject's original fMRI time courses. The transition probabilities, a_{ij}, were set to be equal for all the possible switches. The Baum-Welch algorithm (Rabiner 1989) was employed to solve the problem in equation (1). Details of the algorithm can be found in Scikit Learning (Pedregosa, Varoquaux et al. 2011).

2.2.3 Stability and reproducibility

Different initializations of the Baum-Welch algorithm may lead to convergence into different local optima. To address this issue, we repeated the algorithm 8 times with 8 different initializations on 50 subjects and selected the model parameter, λ, which provided the best objective $L(\lambda; \boldsymbol{O})$ to increase the likelihood of reaching global optimum. To calculate the stability of brain states from different initializations, we

employed the following method on each half of the dataset. We repeated the above procedure 10 times. The result of each run was called a realization of GHMM. An iterative approach similar to the work by Yang *et al.* (Yang, LaConte et al. 2008) was used to match brain states from different realizations. In each iteration, the most similar pair of brain states (with the largest correlation coefficient of Gaussian mean vectors) from different realizations was assigned to the same group. This process is iterated until all the states were assigned or all the remaining correlation coefficients were smaller than 0.8. Note that each realization can contribute at most one brain state to each grouped brain state.

The stability of brain states assessed by summing up all the correlation coefficients between every pair of brain states in the same group (total number of pairs is $10(10-1)/2 = 45$). In our case, the stability for each brain state was the summation of at most 45 correlation coefficients, and the sum was therefore normalized by dividing it by 45. Finally, brain states were sorted in the order of descending stability. The matching and sorting procedure are illustrated schematically in Figure 2A.

In order to test the reproducibility of our model on different datasets, we split the data of 100 subjects into 2 non-overlapping groups of subjects (50 subjects each), and each half of the dataset was analyzed with the GHMM. The brain states extracted from both halves were compared to see whether the brain states were reproducible in different groups of subjects. Based on our results, we were able to find 9 stable and highly reproducible brain states. Therefore, we subsequently set the total number of states to 9 when applying GHMM to resting-state fMRI data.

2.2.4 Brain state sequence decoding

The Viterbi algorithm (Viterbi 1967) was used to decode the optimal brain state sequence, $\{q_t\}, t = 1, 2, \dots, T$, and determine which state the brain is under at each time point for all the scans after obtaining the optimal parameter set $\tilde{\lambda}$ in equation (1). We also calculated the posterior probability score $P(q_t|\boldsymbol{O}, \tilde{\lambda}), t = 1, 2, \dots, T$, of each brain state at all the time points (Rabiner 1989).

2.2.5 Spontaneous brain activity pattern of each state

We decoded the hidden brain state sequence for all the subjects and realizations with the Viterbi algorithm. After matching and grouping the brain states of different realizations, a time frame was labeled with a state only when it was assigned to the same state in all the realizations. Then the z-scores of all the time frames that have been under the same state were computed as a representation of the spontaneous brain activity pattern of the brain state.

2.3 Results

After grouping the brain states from all 10 realizations, the stability of brain states was calculated. The states were subsequently sorted in the order of descending stability (S1 is defined to be the most stable state). Figure 2B and Figure 2C show the stability of the resultant brain states when the total number of states, M, was set to 9 and 10, respectively. When $M = 9$, each of the 9 states was able to match to a corresponding state in all 10 realizations, and thus only 9 groups emerged (Figure 2B). However, when $M = 10$, some states in some realizations failed to find a match in other realizations, so

there were 11 possible groups for this case (Figure 2C). Figure 2B and Figure 2C illustrate that the stabilities of brain states are close to 1 when the total number of states is 9, while some states start to exhibit low stabilities when the total number of states is set to 10. In fact, with varying number of states ($M = 5, 8, 9, 10, 11, 12, 20$), it was found that the stabilities are almost 1 when M is set to 9 or less but substantially less than 1 when M is greater than 9. Therefore, for the subsequent analysis, the number of states was set to 9.

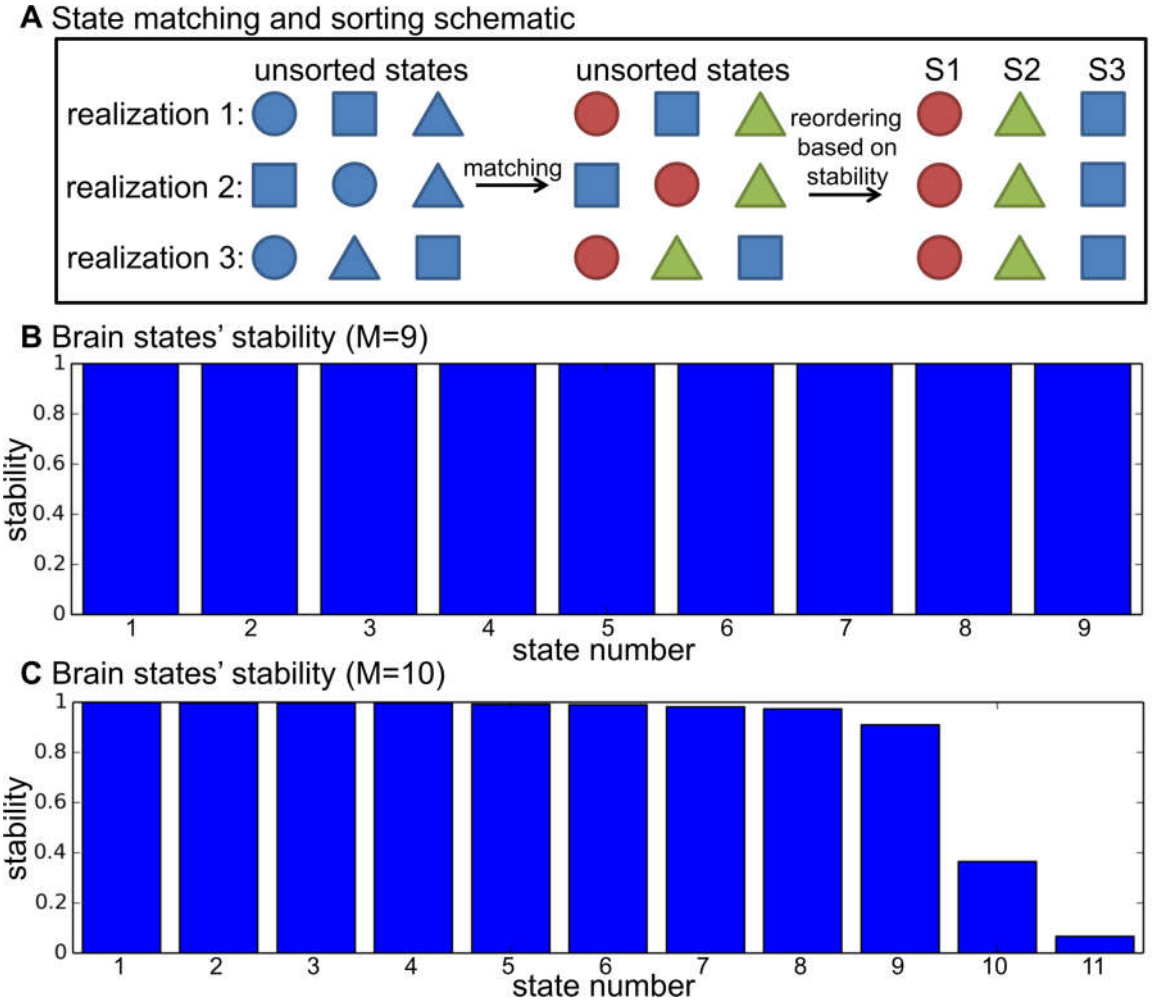

Figure 2 – The schematic of state matching and sorting (A) and the resulting stability of each brain state when the total number of states is set to 9 (B), 10 (C). A toy example of the schematic of the state matching and sorting method with three states and three realizations are shown in (A). Spatial patterns of the brain states and the matched groups are represented by shapes and colors, respectively. The stability of brain states for M=9

and M=10 is shown in (B) and (C). Note that when M=9, brain states from all the realizations are matched perfectly into 9 groups, while when M=10, they are not matched perfectly but grouped into 11 groups instead.

The spatial patterns of the 9 brain states detected from the entire dataset, denoted as S1 to S9, are further demonstrated in Figure 3. For better visualization, Figure 3 only displays z-scores below -20 or above 20. For each state, a brain region with large positive or negative z-scores indicates that it is highly activated or deactivated. Note that the absolute z-scores in S9 are below 20 so no activation is seen in Figure 3 for S9.

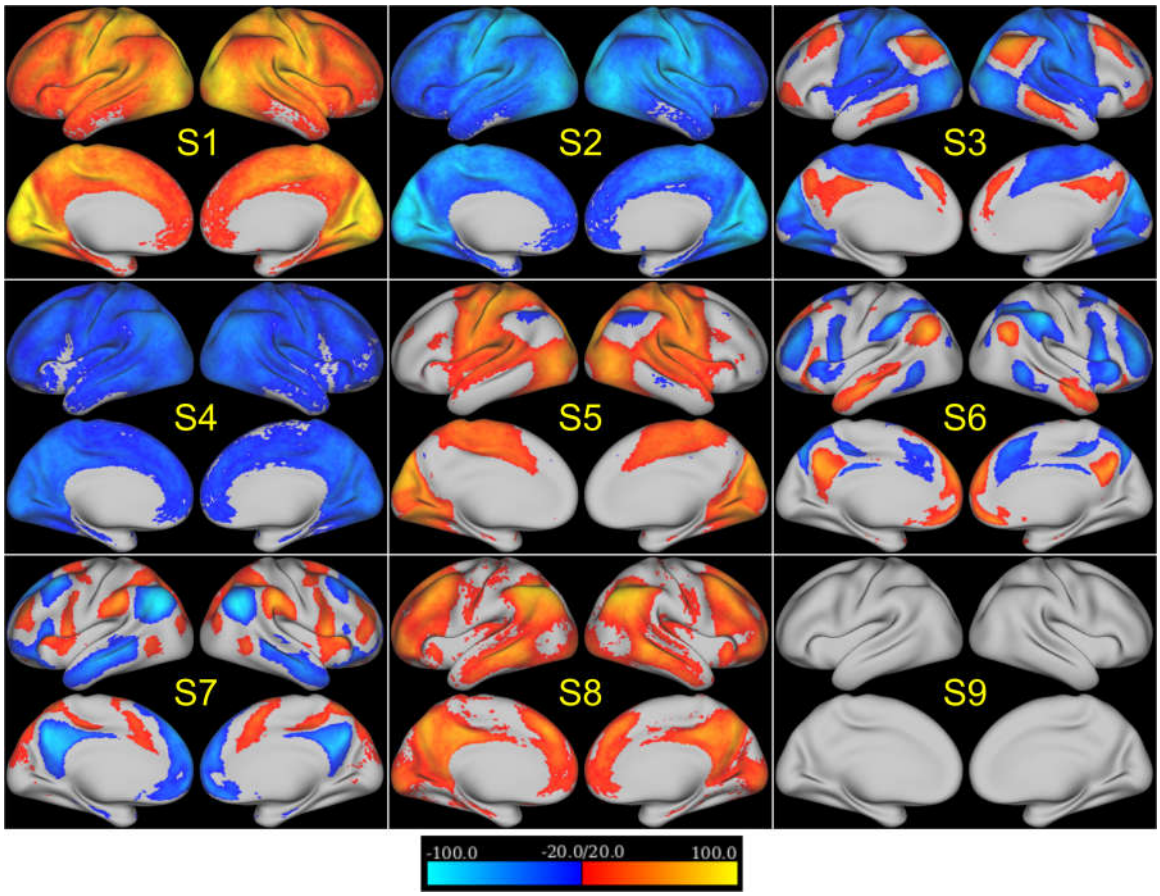

Figure 3 – The z-maps of 9 reproducible brain states. The color bar is set to only display z-scores below -20 or above 20. These brain states are different combinations of activated and deactivated brain regions.

All detected brain states consist of combinations of activated and deactivated brain regions. For example, as shown in Figure 3, brain activities in S6 and S7 are concentrated in regions specific to default mode and attention networks. S6 contains activation in the DMN with deactivation in the attention network, while S7 is comprised of deactivation in the DMN and activation in the attention network. Brain states S3 and S5 are almost opposite in a sense that S5 is comprised of activation in the sensorimotor and visual networks along with deactivation in the frontoparietal control network, while S3 contains deactivation in the sensorimotor and visual networks plus activation in the frontoparietal control network and DMN. Meanwhile, S1 and S2 are whole brain activation or deactivation with sensorimotor and visual networks being the most activated or deactivated. Also, S8 and S4 show whole brain activation and deactivation, but with lower intensities than those in S1 and S2. The final state, S9, has less apparent activation or deactivation patterns in known networks compared with other states.

We also found that the brain is constantly switching among different states. Figure 4A shows an example of a sequence of brain states and how the posterior probability, $P(q_t|O_t, \tilde{\lambda})$, of each state changes over time. In Figure 4A, the state sequence decoded by the Viterbi algorithm is represented by different colors, while the posterior probability is plotted with black curves. Figure 4B illustrates the occurrence and average duration of each state. Note that S9 occurs most frequently while S1 and S2 occur less frequently.

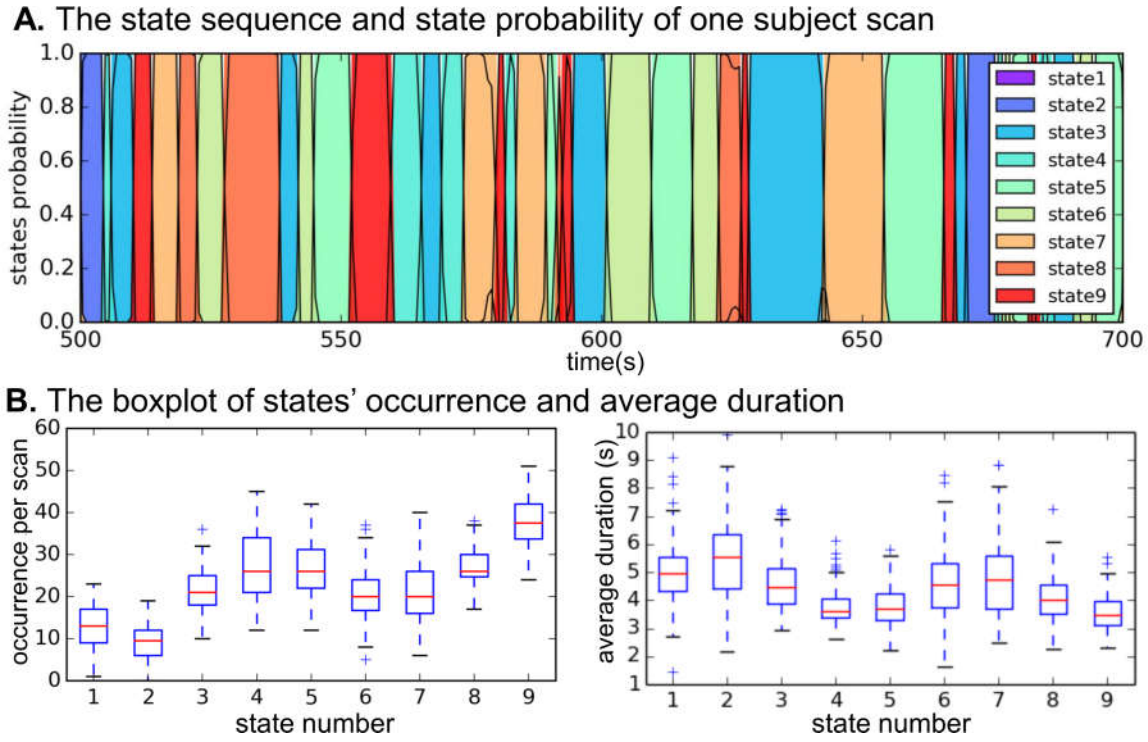

Figure 4 – Temporal characteristics of the brain state. (A) The state sequence and state probability during one subject scan. Only a period of the scan is demonstrated in the figure for better visualization. Different colors indicate what the subject's brain state is at each time point. The black curves are the posterior probability of different brain states. (B) Boxplots of states' occurrence and average duration. Note that the length of each scan is 14.4 minutes.

After decoding the state sequence, we counted the number of transitions between states and calculated the average direct transition times between each pair of states. Only transitions that are significantly more frequent than the average are color-coded in Figure 5A ($p < 0.0007$. Since this is a multiple comparison, Bonferroni correction has been applied to the α value). Note that most of the states (except S1 and S2) transit to S9 frequently, and S9 also switches back to other states frequently.

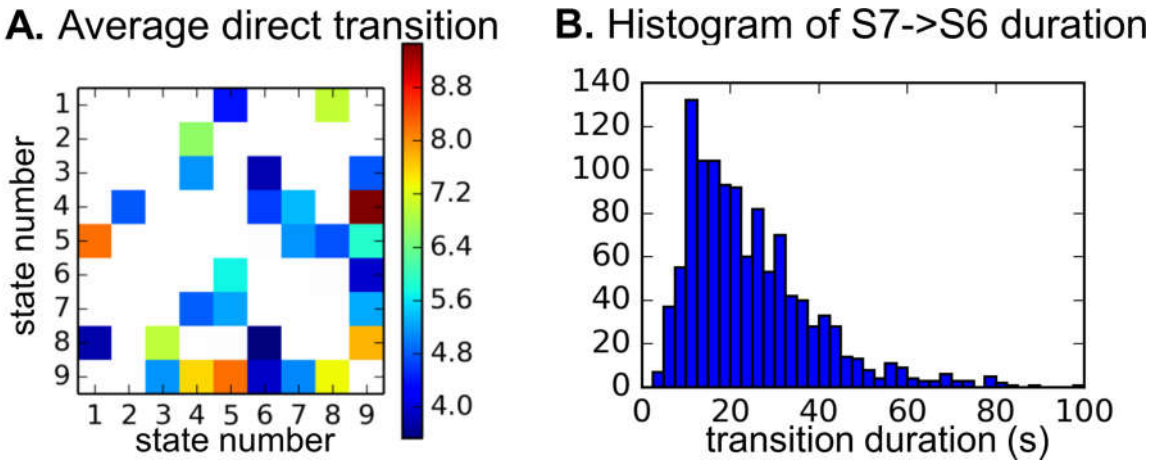

Figure 5 – The transition properties of the brain states. (A) Average direct transition per scan. Note that only when the transition times from state *i* to state *j* is significantly higher than average ($p<0.0007$), the block in the *i*th row and *j*th column will be color coded. (B) Histogram of the duration of the motif from the beginning of S7 to the end of S6.

The transition from the DMN activation to the attention network activation has been noted previously (Majeed, Magnuson et al. 2011). Based on our results, S7 shows attention network activation while S6 exhibits DMN activation. Therefore, we investigated the duration of the motif from the beginning of S7 to the end of S6 (direct and indirect transitions have both been accounted for). Figure 5B shows the histogram of the duration, in which the peak of the distribution lies between 10s and 20s.

To test the reproducibility of brain states, the spatial pattern of brain states detected from the non-overlapping split-half samples of subjects are presented in Figure 6. The 9 brain states extracted from the entire group of 100 subjects and two non-overlapping samples of 50 subjects are almost identical (Figure 6A). As can be seen in Figure 6B, correlation coefficients between the brain states from the two halves of the dataset indicate that the spatial patterns of all brain states are highly reproducible across different groups of subjects ($p \ll 0.001$). Although the correlation coefficient of the last

state is the lowest, it is still statistically significant ($p \ll 0.001$). The last state also has relatively low z-scores (darker colors in Figure 6A and no patterns after thresholding in Figure 3) compared to the other 8 states. The temporal characteristics for these 9 brain states are also tested on split-half samples as shown in Figure 7. These temporal characteristics include the percentage of time occupied by each state, the occurrence per scan, and the average duration of each state. Note that, among all the 27 temporal features of the brain states, only the occurrence of S8 shows a significant difference between two halves (p = 000001). The rest of the temporal features are highly reproducible (p > 0.0019 = 0.05/27).

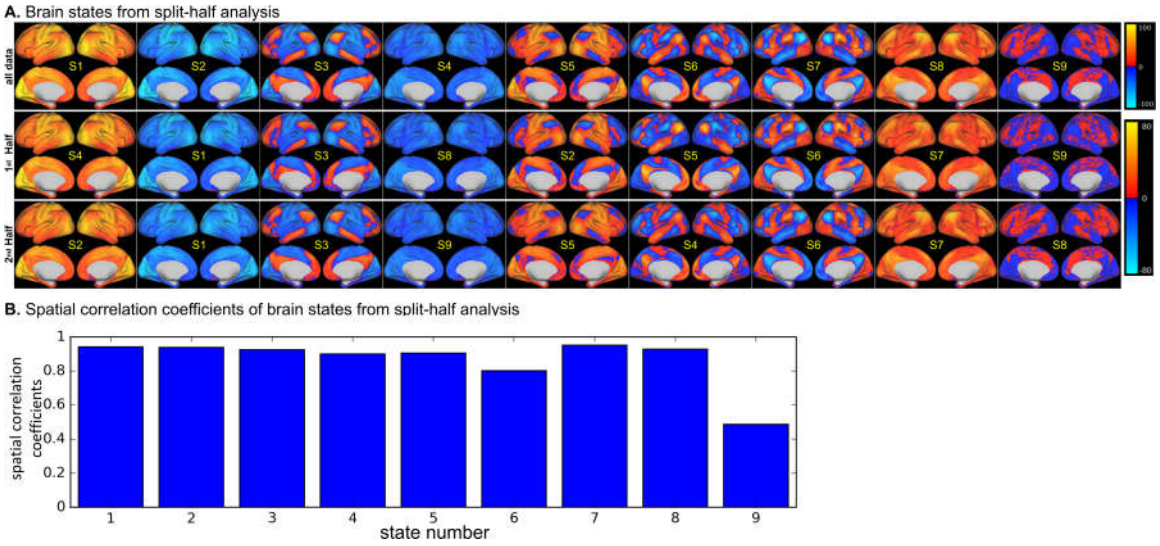

Figure 6 – The reproducibility of brain states for split-half samples. (A) The z-map of brain states detected from all 100 subjects and two non-overlapping halves of the dataset. States with similar spatial patterns are aligned in the same column. (B) The spatial correlation coefficients of two states from the two halves are plotted to show the similarity between them. Note that the last state has the lowest spatial similarity and the lowest z-score on its z-map.

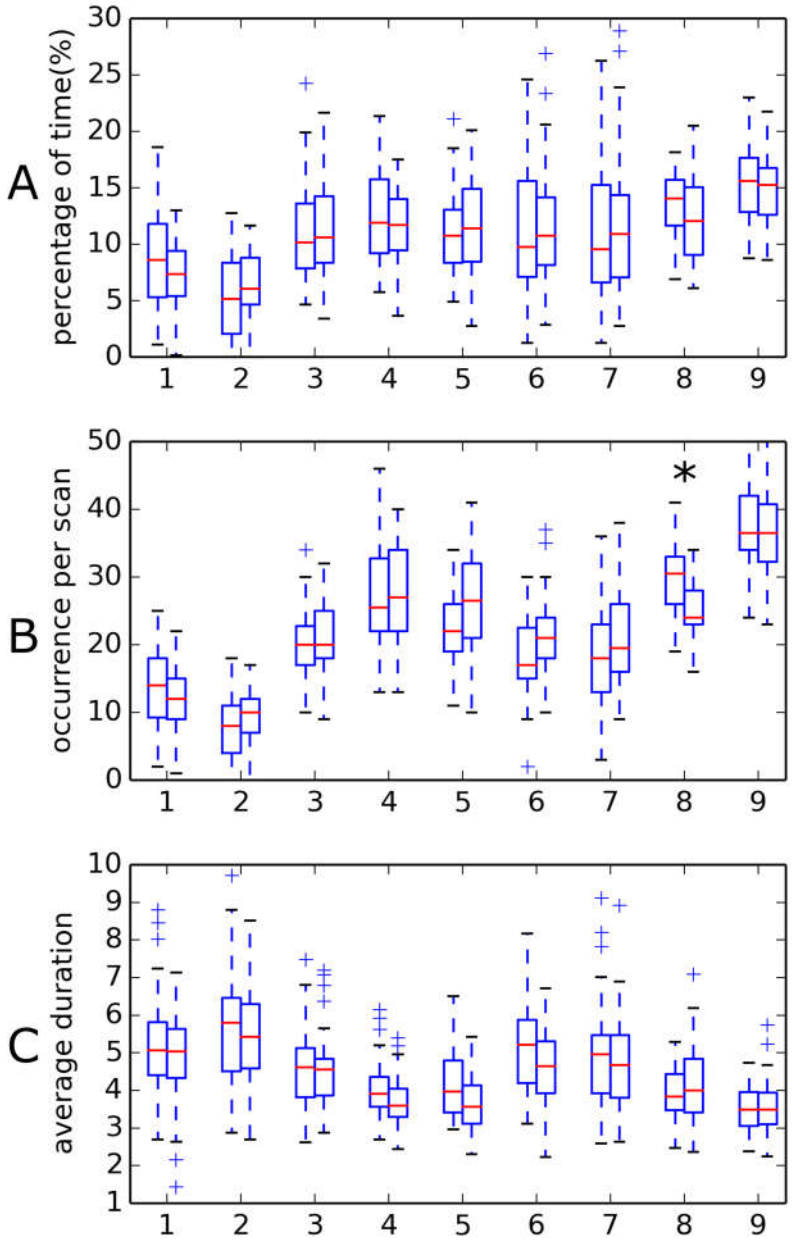

Figure 7 – The reproducibility of the temporal characteristics of brain states for split-half samples. (A) The boxplot of time occupation from two halves; (B) the boxplot of occurrence from two halves; (C) the boxplot of average duration from two halves. The star marks the significant difference between two halves.

To illustrate the advantage of exploiting temporal information, the method used in CAPs (Liu, Chang et al. 2013, Liu and Duyn 2013), k-means, was applied on the signal from the 236 ROIs used in the GHMM analysis. For this analysis, the split-half approach

described in the methods section was repeated and the reproducibility of the states derived with the repeated application of k-means was tested. For comparison, the number of initializations and repetitions used was identical to that for the GHMM analysis. Figure 8 shows the spatial patterns of resultant cluster centers, which are denoted by C1 to C9. Note that for most of the states derived by GHMM (S1, S2, S3, S5, S6 and S8), we can find very similar patterns in these cluster centers derived by k-means, while whole brain deactivation (S4) and attention networks (S7), and their corresponding cluster centers (C1 and C8) have relatively low spatial correlation coefficients. No spatial patterns similar to S9 can be found in these cluster centers (Figure 8C and Figure 6). Moreover, the reproducibility of C5 and C7 are relatively low compared to the rest (Figure 8B).

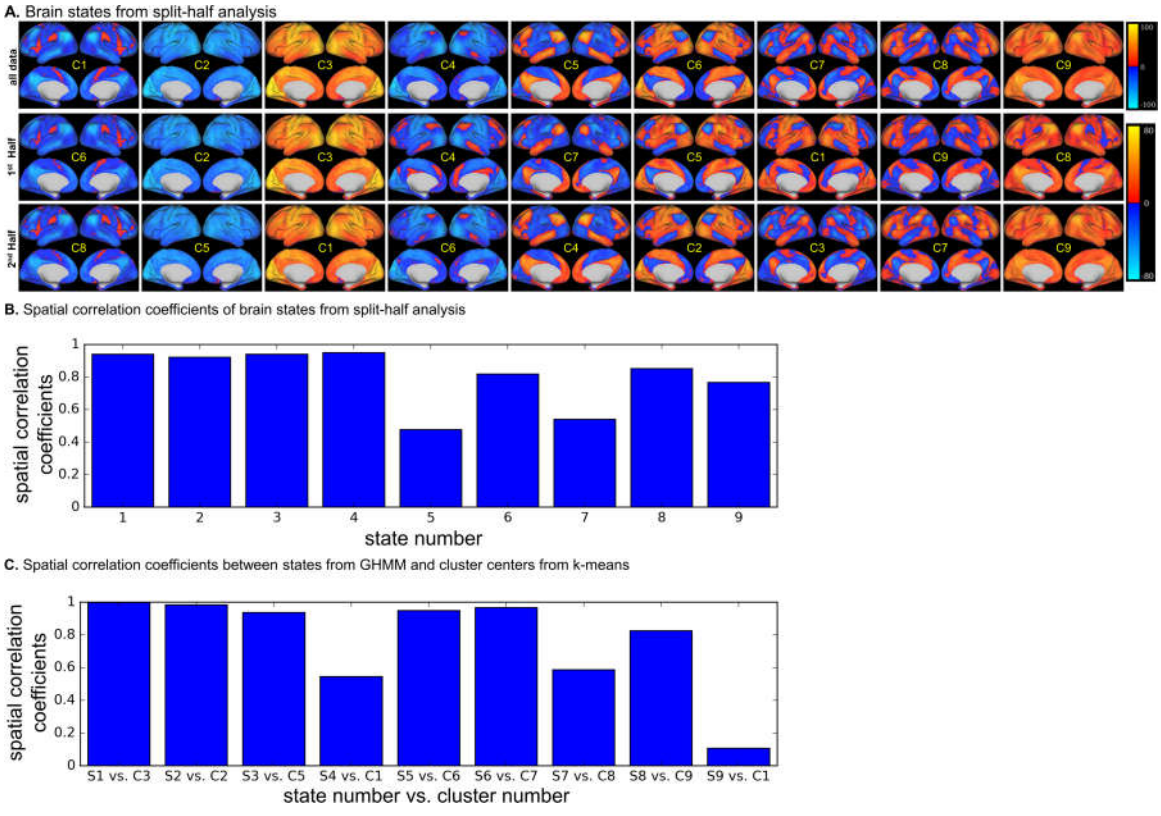

Figure 8 – The reproducibility of cluster centers derived by k-means. (A) The z-map of cluster centers detected from all 100 subjects and two non-overlapping halves of the

dataset. Similar spatial patterns are aligned in the same column. (B) The spatial correlation coefficients of two cluster centers from the two halves are plotted to show the similarity between them. (C) Spatial correlation coefficients between states derived by GHMM and cluster centers derived by k-means.

2.4 Discussion

In general, the Markov property can be applied on either temporal or spatial ax of data. When it is used to model spatial relationship among voxels, it is called a Markov random field and a binary state switch is typically assumed for each voxel (Lindquist, Waugh et al. 2007, Robinson, Wager et al. 2010, Liu, Awate et al. 2014). When the Markov property is assumed on the temporal axis, it is a Markov random process and has been employed to investigate variability in functional connectivity over time (Eavani, Satterthwaite et al. 2013, Ou, Xie et al. 2014). In contrast to the previous application of HMM, we applied the GHMM on temporal axis to model the spontaneous brain activity state switching process. Moreover, we introduced a robust approach for determining the number of states based on the stability. The non-stationary assumption of the GHMM provides the ability to capture the dynamics of the brain activity measured by the resting-state fMRI. Our spatiotemporal model of spontaneous activities across the brain in resting-state fMRI was able to identify 9 stable brain states with reproducibility near 1. Furthermore, the detected brain states have analogs to combinations of conventional resting state networks (RSNs). The implications of these results are discussed in the following paragraphs.

The reproducibility of GHMM on estimating brain states is affected by two factors: A) algorithmic stability and B) the dataset that the model is trained on. Since the algorithm is non-convex, there is no guarantee of a global optimum and different

initializations can lead to different local optima. Increasing the repetitions of the algorithm with different initializations will increase the chance for the algorithm to reach the global optimum and, thus, increase the stability of the results. Increasing the total number of states will increase the number of parameters in the model which may lead to a more complicated model with more local optima and thus a less stable result. This could explain the low stability of some states when the total number of states is larger than 9 (Figure 2C). Meanwhile, different datasets, O, will result in different objective functions, $P(O|\lambda)$, and thus different best fitting model parameters λ. This is due to the fact that subjects in different fMRI scans may experience different brain states. Our results (Figure 2 and Figure 6) show that brain states detected by GHMM are both stable and reproducible on different datasets, indicating that the model is able to extract the common features within this group of subjects. The stabilities of all 9 states are almost 1, as shown in Fig 2B, indicating the brain states from all realizations are very similar and can be matched perfectly into 9 groups. Therefore, changing the threshold, 0.8, to a higher value, when matching states, will not lead to any differences in the final results. Moreover, comparing the reproducibility of the cluster centers derived by k-means and our model (Figure 8B), the higher reproducibility of our result (Figure 6B and Figure 7) is due to additional temporal information included in our model.

No spatial constraint is imposed in our GHMM, yet the 9 states identified exhibit smooth contiguous activated and deactivated regions. We further note that these activated and deactivated brain regions resemble conventional RSNs. For example, the whole brain deactivation states, S2 and S4, have been found by temporal functional modes (TFMs) (Smith, Miller et al. 2012). Also, linear combinations of activated/deactivated DMN and

deactivated/activated attention network, S6 and S7 (See Figure 3), have also been reported as CAPs (Liu, Chang et al. 2013, Liu and Duyn 2013). The existence of these two brain states explains why DMN and attention network are often anti-correlated in seed-based correlation analysis (Fox, Corbetta et al. 2006). Our results further indicate that these two networks are only anti-correlated when the brain is in S6 and S7, rather than during the entire scan. Moreover, the switch between different states leads to changing correlation and anti-correlation patterns and, thus, can explain why the functional connectivity changes over time (Allen, Damaraju et al. 2014). Comparing Figure 6 and Figure 8A, we note that most of the spatial patterns of brain states detected by our GHMM (S1, S2, S3, S5, S6, and S8) are similar to those detected by k-means (Figure 8C), indicating most of our brain states are consistent with those detected by CAPs.

As shown in Figure 4, the duration of states differs for each state and even differs for each occurrence. Therefore, the use of conventional sliding-window-based analysis with fixed window length can lead to signals from multiple brain states being merged together and possibly canceling each other. Based on our results, we recommend decoding the brain state sequence first and then determining the window length according to the duration of states rather than fixing the length and sliding the window.

To the best of our knowledge, a state similar to S9 has not been presented in the previous literature. We hypothesize that S9 is the "ground" state of the brain, in which brain activity (or deactivity) are similar for the entire cortex (no apparent activation or deactivation as shown in Figure 3). Note that different groups of subjects have different spatial patterns for state S9 (Figure 6A). Therefore, S9 has the lowest reproducible spatial

pattern (Figure 6B). However, its temporal characteristics allowed us to distinguish it consistently from other states. S9 occurs more frequently (37.5±6.4 times per scan. or per 14.4 minutes) compared to other brain states and has a short duration (3.6±0.6 s) (Fig 5B). From the transition matrix in Figure 5A, we can see that most states (except S1 and S2) often switch to S9 and then transit back to other states from S9. The spatial and temporal characteristics of S9 indicate that it is an intermediate transient state that appears when the brain is switching between other more reproducible brain states and could represent a "ground" state. Note that the low reproducibility of S9's spatial pattern makes it difficult to detect using methods that only consider the spatial information when defining states. This may be the reason why k-means cannot detect this state (Figure 8A). Our ability to detect it could be due to the additional temporal information inherent in our model.

States S7 and S6 have similar spatial activation patterns as the attention network and DMN, respectively, reported in the literature (Buckner, Andrews‐Hanna et al. 2008, Broyd, Demanuele et al. 2009) (Figure 9). Prior work (Majeed, Magnuson et al. 2011) found a reproducible transition from activation in attention network (similar to S7) to activation in DMN (similar to S6) within a 20s sliding window. Interestingly, the present method also detected similar transition duration between these states. The histogram of the motif from the beginning of S7 to the end of S6 in Figure 5B shows that there exists a peak between 10s and 20s, indicating that the motif is more likely 10s to 20s in duration, consistent with what was reported in the literature (Majeed, Magnuson et al. 2011). Note that this transition from S7 to S6 not only accounts for direct transition but also includes

indirect transitions via intermediate states (e.g. S7->S9->S6), which was also shown by Majeed et al. (Majeed, Magnuson et al. 2011).

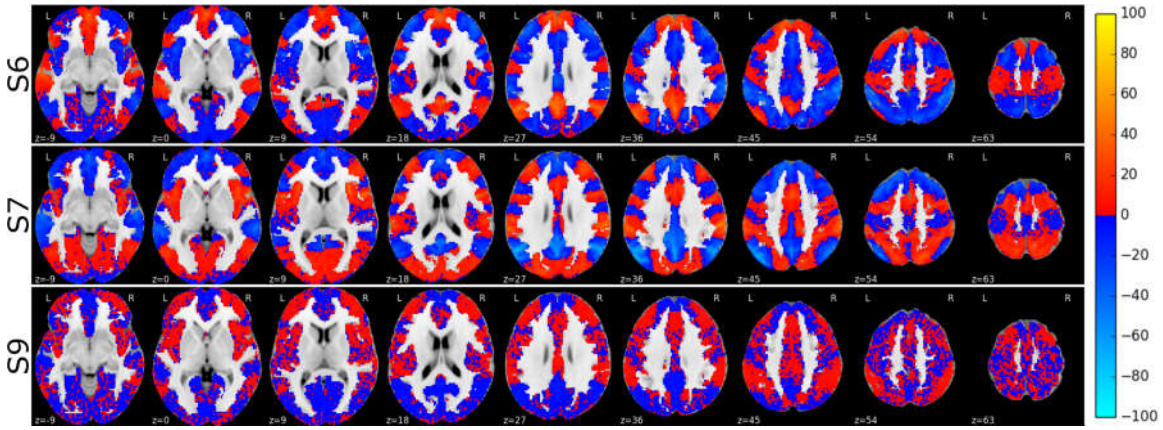

Figure 9 – Whole brain axial image for S6, S7, and S9.

The intrinsic advantage of GHMM over CAPs is that GHMM quantifies state transitions explicitly using a Markov chain in the model. The quantification of state transition not only provides the ability to capture temporal characteristics (Figure 4 and Figure 5) but also give rise to additional brain states. Specifically, the 9th state, detected by GHMM, due to a relatively low reproducibility of its spatial pattern, is difficult to capture by methods that only consider spatial similarity, such as k-means used by CAPs. However, since S9 has distinctive temporal properties (high occurrence frequency and short duration), GHMM is able to detect it consistently (Figure 6). Figure 8C further illustrates that there are spatial differences in activation patterns between CAPs and GHMM in other brain states as well, including the whole brain deactivation (C1 for CAPs and S4 for GHMM) and attention networks (C8 for CAPs and S7 for GHMM).

The results in this work must be considered in the context of several limitations. First, we assume a memory-less transition between brain states by using Markov chains,

i.e., the current brain state only depends on the previous state. Fortunately, the HMM is robust to violations of this assumption and has been successfully applied to speech recognition (Rabiner 1989) as well as detection of brain states from electrophysiological measurements (Jones, Fontanini et al. 2007), neither of which is a memory-less system. Another methodological limitation is that state duration is not modeled explicitly in the present work, in order to stay in the same state, the system needs to conduct a self-transition, which would decrease the probability of staying in the same state exponentially over time. In future work, the HMM with explicit duration can be applied, and the duration probability distribution will be trained with the dataset as well. Due to computational limitations, we only trained our GHMM with 236 time courses from the brain of each subject. A voxel-wise GHMM may be able to capture more detailed spatial characteristics of the states. In present work, we have applied temporal filtering (0.01–0.1 Hz) to focus on the low frequency spontaneous activity (Fox, Snyder et al. 2005). However, recent studies have demonstrated that visual and sensorimotor networks are also present in higher frequencies (0.5-0.8Hz) (Lee, Zahneisen et al. 2013). We plan to study the effect of different temporal resolutions by applying GHMM to high temporal resolution data and study brain state switching process at other time scales.

2.5 Conclusion

We have applied GHMM to analyze resting-state fMRI. The GHMM models the brain state switching process as a Markov chain and the brain states as multivariate Gaussian distribution. Spatially, 9 stable and reproducible brain states were discovered and found to be combinations of activated or deactivated RSNs. Temporally, we were able to derive the brain state sequence for individual subjects with the model and evaluate

the occurrence and duration of each state. One transient state, S9, was identified based on its spatiotemporal characteristics. The motif from activated attention network (S7) to activated DMN (S6) was also found to be consistent with previous literature. Therefore, we conclude that the study of brain state sequential switching process can unveil spatiotemporal pattern of the brain and further improve our understanding of brain dynamics.

CHAPTER 3. REPRODUCIBILITY OF BRAIN STATES DETECTED BY GHMM OF AND ITS APPLICATIONS

3.1 Introduction

In the previous chapter, GHMM was used to model brain dynamics using fMRI data. We have derived 9 stable and reproducible healthy control brain states using the HCP dataset. We have also demonstrated both the spatial and temporal characteristics of these states. In this chapter, we will test how different datasets and different preprocessing approaches will influence the reproducibility of GHMM. After we ensure the reproducibility of GHMM, we will apply the model on biomarker detection of prenatal alcohol exposure (PAE).

Due to the relatively low frequency hemodynamic response of neural activity, rfMRI analysis often only considers some low frequency fluctuations of the signal (bandpass filtered to include only 0.01 – 0.1Hz). Some researchers have also detected high frequency functional connectivity using rfMRI (Trapp, Vakamudi et al. 2017), which puts the necessity of conducting temporal filtering as a preprocessing step under debate. However, how this high frequency signal in fMRI changes the brain state switch process is still unknown. With the fast acquisition in HCP data, we are able to investigate the effect of wider frequency band fMRI signal.

Whether or not we should regress out the global signal is another controversial topic. The global signal is the average time course of all the voxels in the brain. Physiological signals such as respiration, the cardiac cycle, and head motion have been

shown to be highly correlated with the global signal (Burgess, Kandala et al. 2016). However, the correlation itself does not guarantee that global signal does not have any neuronal component. In fact, some research even shows that GSR can cause artificial anti-correlation between brain regions (Fox, Zhang et al. 2009). In this work, the effect of GSR on brain states will be investigated by comparing the GHMMs before and after GSR.

Other factors are investigated in this chapter include the effect of advanced imaging techniques and differing brain atlases. The emergence of fast acquisition techniques such as multi-band, has allowed fMRI data to be acquired with significantly improved spatial and temporal resolution (Setsompop, Gagoski et al. 2012). How our GHMM will perform on different datasets acquired with and without fast acquisition will be examined. The effect of using different brain atlases will also be assessed. Our GHMM is trained based on time courses extracted from a few ROIs. Some coordinate based atlases are generated based on a meta-analysis of fMRI (Power, Cohen et al. 2011), while some atlases are combinations of multimodal results (Glasser, Coalson et al. 2016). Affects arising from the use of different atlases on the detected brain states will be investigated in this chapter as well.

PAE, the exposure to alcohol during pregnancy, is known to hinder fetal development and also retard post-birth development. The severity of PAE varies based on the timing, duration, and amount of alcohol consumed during pregnancy, as well as maternal health and socioeconomic factors. Individuals who satisfy the clinical criteria for fetal alcohol syndrome (FAS) show deficits in growth, morphology, and central

nervous system development. Individuals with PAE, including those who do not meet criteria of FAS, show cognitive and behavioral deficits (Larkby and Day 1997).

The underlying brain function alterations caused by PAE have been investigated using functional neuroimaging, especially functional magnetic resonance imaging (fMRI) (Coles and Li 2011). Most of these studies require subjects to conduct certain tasks and only examine those task-related brain regions. A wide range of differences in activation patterns between PAE and controls have been reported in the previous literature (Coles and Li 2011). It is difficult to interpret these results because different studies focus on a variety brain regions and some of them are even inconsistent. For example, both (Malisza, Allman et al. 2005) and (Spadoni, Bazinet et al. 2009) focused on working-memory-related brain regions; however, (Malisza, Allman et al. 2005) showed PAE has less brain activity overall while (Spadoni, Bazinet et al. 2009) demonstrated greater response in most brain areas of PAE. This inconsistency is probably due to the differences in task design and subjects' behavioral performance. These works have focused on some specific activation in task-related brain regions.

rfMRI and functional connectivity have also been employed to investigate PAE. For example, Wozniak and colleagues (Wozniak, Mueller et al. 2011) reported a decrease of inter-hemispheric functional connectivity in children with PAE. As another example, Ji and colleagues (Ji, Li et al. 2015) demonstrated differences in independent components between PAE and control groups. These differences have been detected in multiple brain networks, including the default mode network (DMN), left prefrontal/parietal network (LPFN), right prefrontal/parietal network (RPFN), primary motor network (PMN), primary visual network (PVN), and extrastriate visual network (ESVN).

Although previous neuroimaging studies have discovered a wide range of differences in activation patterns and functional connectivity between PAE and controls, all of these studies used stationary approaches and some only focused on specific brain regions or connectivity. Given the broad range of alterations reported in the current literature, we conducted an exploratory study to better characterize PAE by identifying the spatiotemporal characteristics of affected brains, and defining them as disease biomarkers.

3.2 Materials and Methods

3.2.1 Dataset

In order to study the effect of some preprocessing steps, including GSR and the choice of atlases, we used data from 100 healthy individuals of the HCP dataset (age: 22-36+; gender: 46M/54F; TR=0.72s; duration: 14.4minutes). The spatial resolution of the HCP dataset is 2 mm × 2 mm × 2 mm.

To test the reproducibility of brain states detected by GHMM on different datasets, we included another 90 healthy individuals from the Beijing_Zang dataset (age: 18-26; gender: 28M/62F; TR: 2s; duration: 7.5 minutes). The spatial resolution of Beijing_Zang dataset is 3.1 mm × 3.1 mm × 3.6 mm.

Eighty-seven subjects (age: 7-17; gender: 48M/39F, TR: 2, duration: 6 minutes) were recruited to investigate the effect of PAE on brain function. The control group was comprised of 47 subjects of matched age, gender, race, and socioeconomic factors. The spatial resolution of this dataset is 3.4375 mm × 3.4375 mm × 5 mm.

3.2.2 Preprocessing

For the HCP dataset, preprocessing was performed according to the HCP minimal preprocessing pipelines (Glasser, Sotiropoulos et al. 2013). In particular, we considered the data that had been processed with ICA-FIX (Griffanti, Salimi-Khorshidi et al. 2014, Salimi-Khorshidi, Douaud et al. 2014), registered onto 32k Cento69 surface mesh (Van Essen, Glasser et al. 2012) and smoothed with a 2 mm FWHM kernel. All time courses were filtered with a 0.01–0.1 Hz band-pass filter.

The configurable pipeline for the analysis of connectomes (CPAC) (Sikka, Cheung et al. 2014) was employed to preprocess the volumetric fMRI data from Beijing_Zang and the PAE dataset. The pipeline was the successor of the 1000 functional connectomes project preprocessing pipeline (Biswal, Mennes et al. 2010), using AFNI (Cox 1996), FSL (Smith, Jenkinson et al. 2004) and ANTs (Avants, Tustison et al. 2011). fMRI data was head motion corrected, skull-stripped, co-registered to skull-stripped high resolution anatomical data acquired from the same subject, and then normalized to the 3mm 152-brain Montreal Neurological Institute (MNI152) template so that group level analysis could be conducted. The signal from linear, quadratic, white matter, cerebrospinal fluid (CSF) signals, and 6 affine motion parameters were regressed out. Note that we also used compcor with 5 components during regression on the PAE dataset for further denoising (Behzadi, Restom et al. 2007).

GSR was not performed for HCP surface data and Beijing_Zang volumetric data. Instead, a band-pass filter with the same cut-off frequency (0.01–0.1 Hz) was used to reduce physiological noise. Similar to the previous chapter, signals from 236 regions of

interest (ROIs) (Power, Cohen et al. 2011) were extracted by averaging time courses of voxels within a 5 mm radius for both datasets. This distance was calculated based on Dijkstra's distance (Hanke, Halchenko et al. 2009) for HCP data and Euclidean's distance for Beijing_Zang data. The mean value of each time course was subtracted, and the standard deviation was normalized to 1 before these time courses were fed into the GHMM.

We also processed the HCP data with three different preprocessing steps: 1) no GSR and 236 ROIs from (Power, Cohen et al. 2011); 2) with GSR and 236 ROIS; 3) with GSR and 360 ROIs from an atlas derived based on HCP data (Glasser, Coalson et al. 2016). Note that since HCP data is in grayordinate, which includes only cortical surface and subcortical regions, we only included those regions when computing the global signal. This approach, mean grayordinate time series regression, is a proxy for GSR on HCP data (Burgess, Kandala et al. 2016). Note that we did not perform any temporal filtering when comparing different preprocessing steps on HCP data.

3.2.3 Training GHMM

When testing the reproducibility of GHMM and the effect of preprocessing approaches on brain states, one GHMM with 9 states was trained on each dataset with each preprocessing approach. When studying the effect of PAE on brain states, two different training methods were compared. First, we trained two GHMMs on the healthy control and PAE groups separately to demonstrate the effect of PAE on the entire brain state switching system. Since each group only has 40^{+} subjects, we were only able to derive seven stable brain states for the PAE analysis. When two different GHMMs were

trained on two different groups, we were not able to get the correspondence between the brain states across groups. Therefore, one GHMM was trained on all the subjects, and the difference in each brain state between the two groups was investigated to show how PAE affects each brain state.

3.3 Results

3.3.1 Reproducibility of brain states on different datasets

Figure 10 shows the spatial patterns of brain states derived from the HCP and Beijing_Zang datasets. Note that the brain states from the Beijing_Zang dataset (Figure 10B) are sorted to match the order of the brain states from the HCP dataset (Figure 10A). Since the two datasets had different spatial resolutions and different data formats, it was not possible to compute the similarity score as we did when estimating the reproducibility of different subject groups within one dataset. However, we noticed that most of the spatial patterns of brain states are similar across different datasets. Among all the brain states, the ground state (S9) shows the largest spatial difference across datasets.

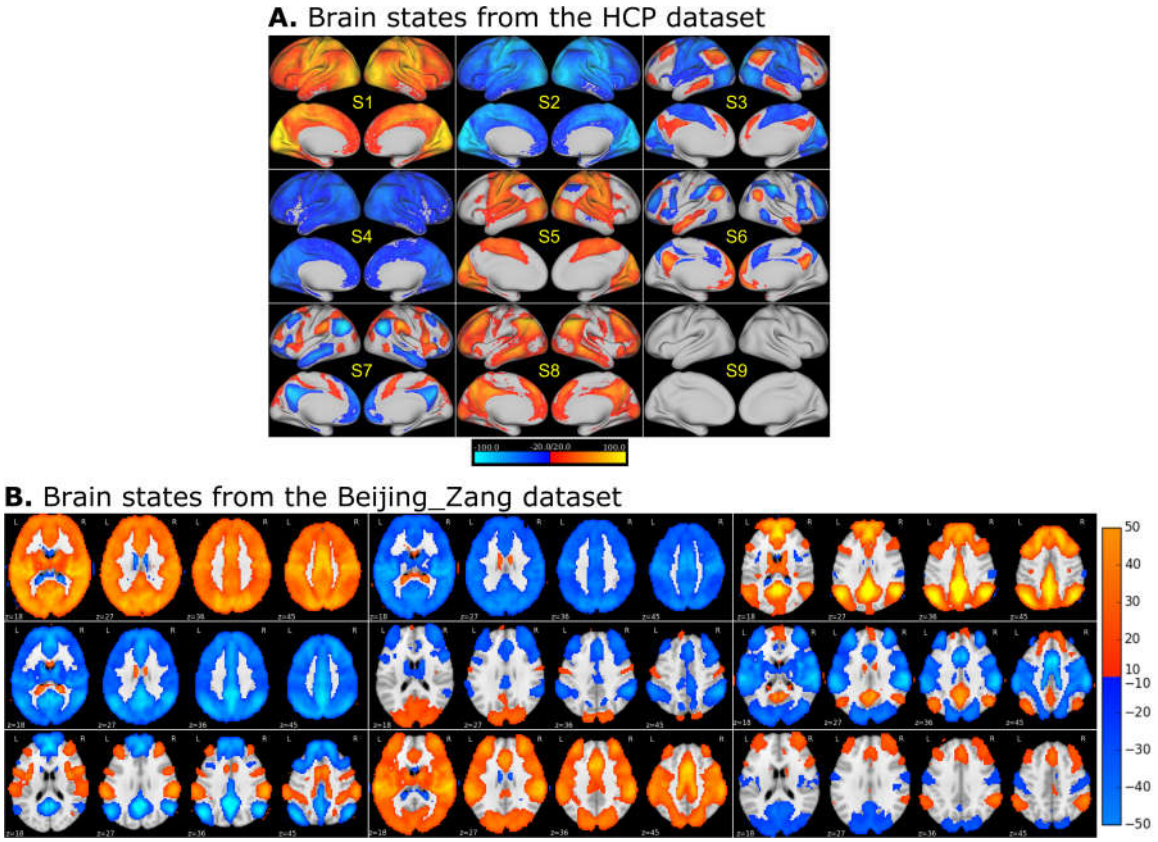

Figure 10 – The reproducibility of brain states' spatial patterns. Z-maps of (A) the brain states derived based on the HCP dataset and (B) the brain states derived based on the Beijing_Zang dataset.

We have also tested the reproducibility of the temporal characteristics of the brain states in Figure 11. These temporal characteristics include the percentage of time occupied by each state, the occurrences per minute, and the average duration of each state. It is worth noting that some of the 27 brain state temporal features start to show a significant difference between datasets. For statistical test, we used the Wilcoxon rank-sum test. The time occupation of S2 and S3 are significantly different with p values of 6.4×10^{-4} and 7.2×10^{-4} respectively. The occurrence of S3, S8, and S9 shows a significant difference between the two datasets with p values of 4.6×10^{-7}, 1.7×10^{-7}, and 1.2×10^{-6}, The average duration of S2, S5, and S9 are significantly different with p values of

3.7×10^{-5}, 1.2×10^{-3}, and 1.2×10^{-3}. The rest of the temporal features are highly reproducible ($p > 0.0019 = 0.05/27$).

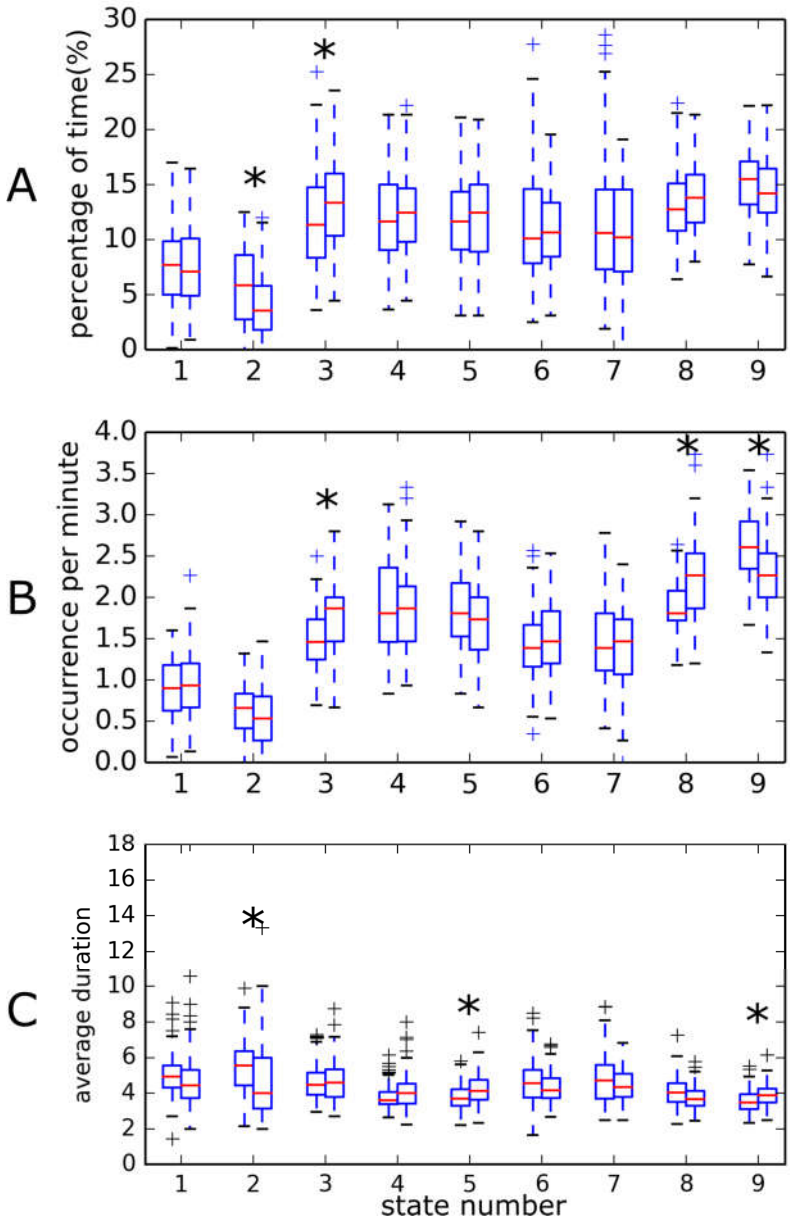

Figure 11 – The reproducibility of the temporal characteristics of brain states between two datasets (Left: HCP; right: Beijing_Zang). (A) The boxplot of time occupation; (B) the boxplot of occurrence; (C) the boxplot of average duration. The star marks the significant difference between two datasets.

3.3.2 Brain states under different preprocessing steps

Temporal filtering (0.01–0.1Hz) is often used to remove unwanted temporal signal in fMRI data. We compared the derived brain states before and after temporal filtering in Figure 12. Note that having the same state number in this figure does not imply they are the same brain state before and after temporal filtering. The states were organized for visualization and comparison purposes. Whole brain activation and deactivation patterns appear in S8 and S9 both before and after temporal filtering. Positive default mode network (DMN) plus negative task positive network (TPN) and its opposite patterns (S1 and S2) are also similar with and without temporal filtering. S6 and S7 with activation/deactivation in the somatosensory motor network (SMN) and visual network (VN) appear regardless of temporal filtering, although the activated region in S7 is part of fronto-parietal network (FPN) before filtering and is more similar to DMN after filtering. Before temporal filtering, S4 and S5 show an activated/deactivated DMN network with some slight deactivation/activation in cingulo-opercular network (CON) region, while after temporal filtering S4 and S5 show whole brain activation/deactivation patterns with DMN being the most activated/deactivated region. The temporal filtering causes great differences to the rest of the brain states. After temporal filtering, a "ground" state appears as described in the previous chapter, which is S3 in Figure 12B with no apparent activation or deactivation patterns. In contrast, we have positive TPN and VN plus negative DMN before temporal filtering (S3 in Figure 12A).

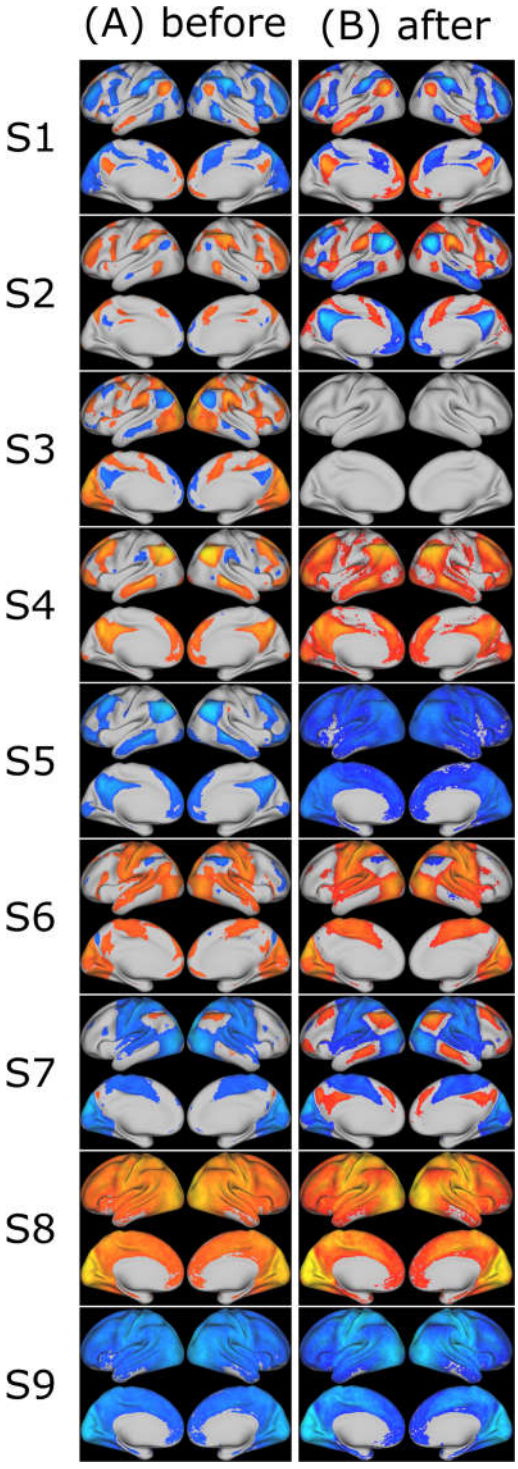

Figure 12 – The spatial patterns of brain states before (A) and after (B) temporal filtering (0.01-0.1Hz). Note that no GSR was performed and 236 ROIs were used to extract each time series. The state number here does not imply they have one-to-one correspondence before and after temporal filtering. It is only for visualization and comparison purposes.

The temporal features of brain states are also significantly affected by temporal filtering. In Figure 13, 3 different temporal characteristics (time occupation, occurrence, and average duration) of the brain states before and after temporal filtering are shown in 6 images. Since two different GHMMs were trained before and after temporal filtering, there is no clear correspondence between the resultant brain states. We plotted their temporal features side by side. The occurrence of all the brain states drops significantly after temporal filtering, while the state duration increases significantly as a result of temporal filtering ($p \ll 0.001$).

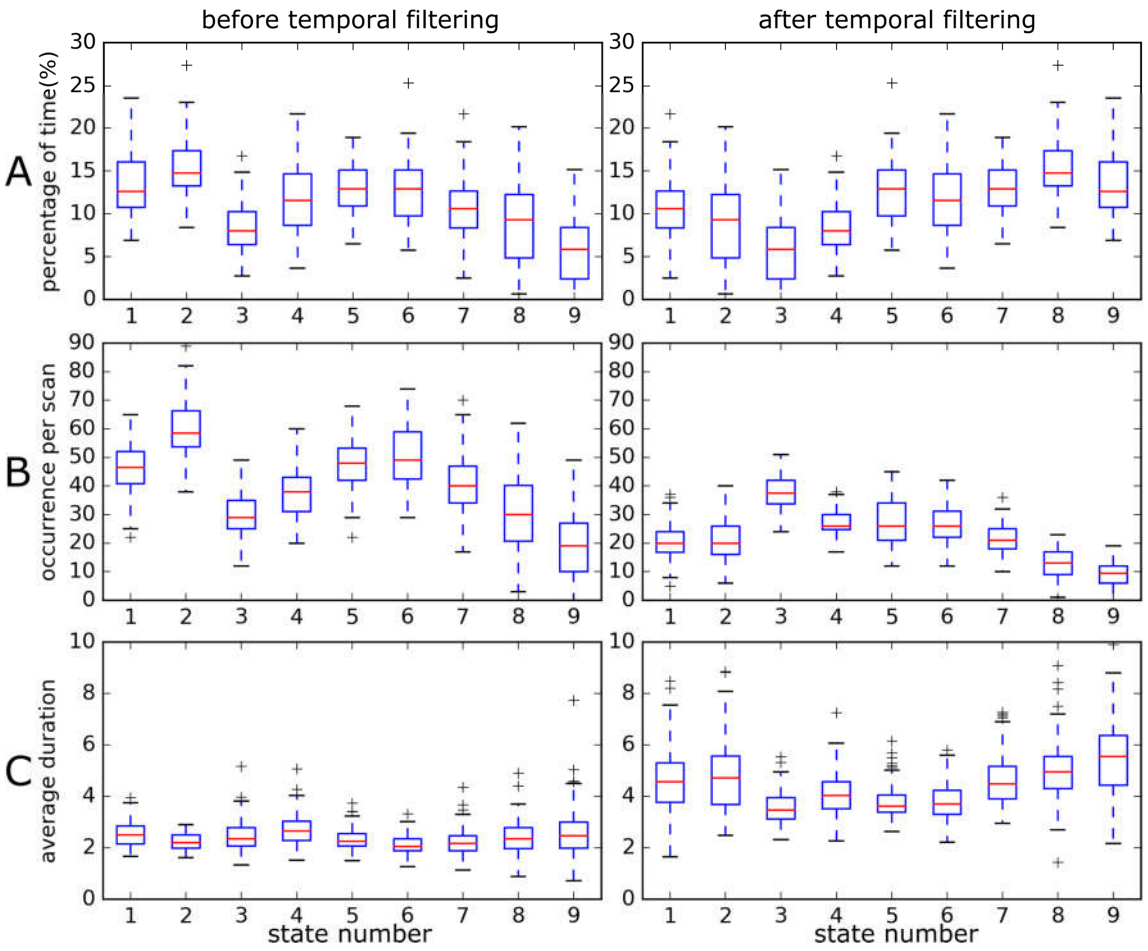

Figure 13 – The temporal characteristics between the corresponding brain states derived before (left) and after (right) temporal filtering. The temporal characteristics

are compared using boxplots of (A) time occupation, (B) occurrence, and (C) average duration.

Figure 14 exhibits the spatial patterns of brain states derived under three different preprocessing approaches. By comparing Figure 14A and Figure 14B, we notice differences in almost every brain state caused by GSR. In Figure 14A, S1 is activation of DMN and deactivation of TPN, while S2 is the opposite of S1. However, in Figure 14B, S1 is a combination of DMN activation and FPN deactivation, while S2 is its opposite. Note that TPN consists of three sub-networks including FPN, dorsal attention network (DAN), and CON. S3 in Figure 14A is activation in visual and CON plus deactivation in DMN, while in Figure 14B, S3 is composed of activation of CON and deactivation of DMN. S4 and S5 in Figure 14A are activation and deactivation of DMN, while S4 and S5 in Figure 14B also show deactivation and activation of visual and TPN. S6 and S7 of Figure 14A and Figure 14B are similar. They both show a combination of FPN and SMN plus VN. Lastly, S8 and S9 in Figure 14A are whole brain activation and deactivation, while S8 and S9 in Figure 14B show activation and deactivation only in visual areas.

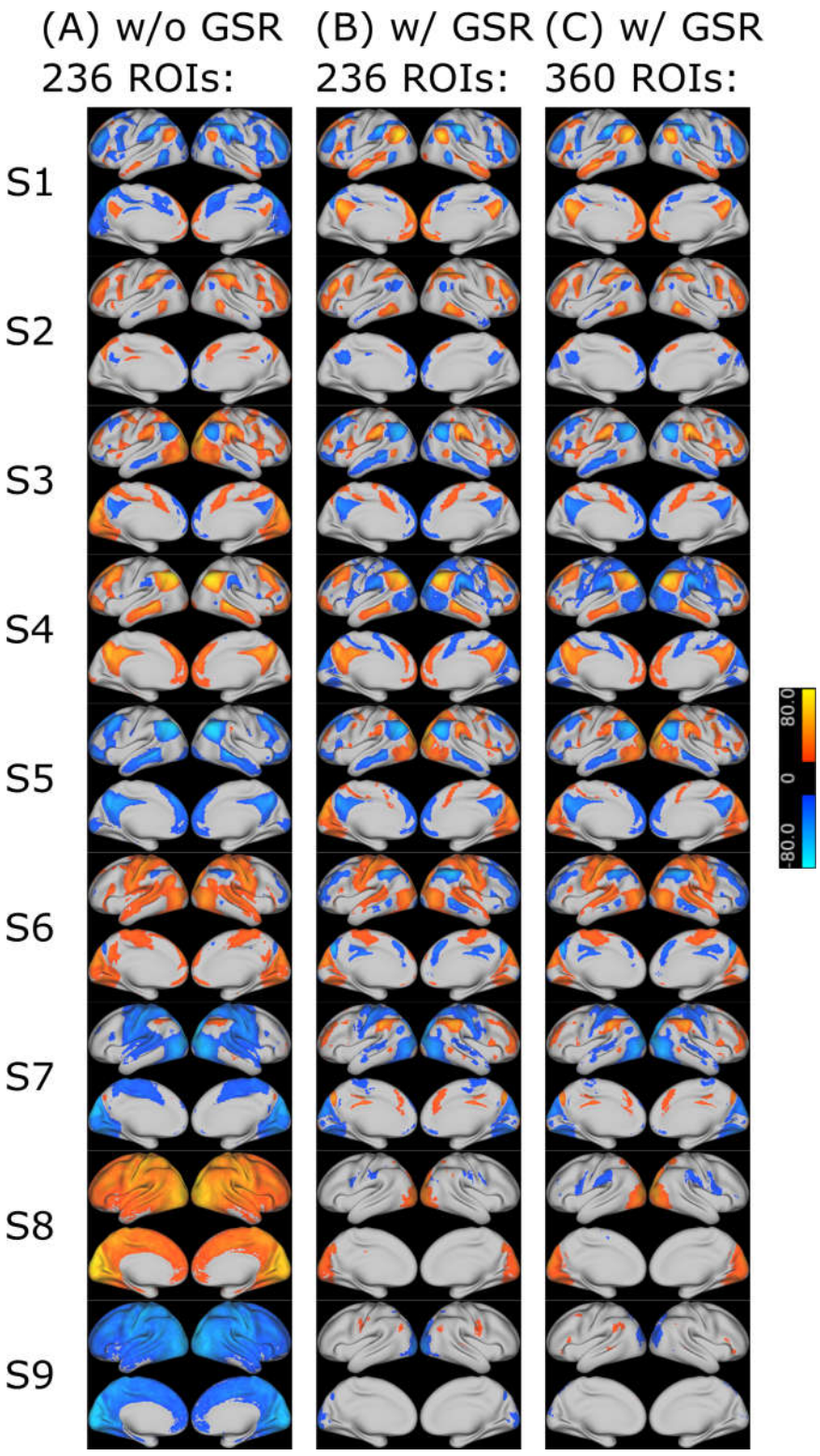

Figure 14 – The spatial patterns of brain states under three different preprocessing approaches. (A) Data extracted from atlas with 236 ROIs, without GSR; (B) atlas with 236 ROIs, with GSR; (C) atlas with 360 ROIs, with GSR.

When comparing the brain states derived with 236 ROIs and 360 ROIs, the spatial patterns are almost identical (Figure 14B and Figure 14C). Figure 15 demonstrates the spatial correlation coefficients between corresponding brain states derived based on 236 ROIs and 360 ROIs with GSR. All of the brain states show significant spatial similarity between two different atlases ($p \ll 0.001$). However, some of the temporal features of these brain states are changed significantly by using a different atlas. In Figure 16, the percentage of time occupation of S3 is changed significantly; the occurrence of S1, S2, S3, S4, and S9 are increased significantly as the number of ROIs increases from 236 to 360. The average duration of all the states decreases significantly when the number of ROIs increases.

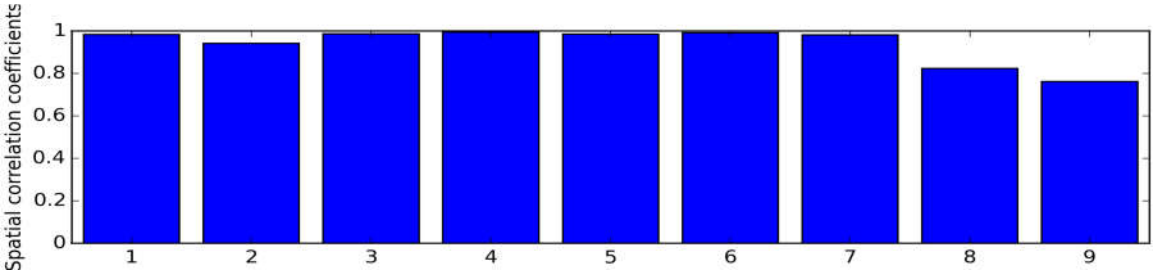

Figure 15 – The spatial similarity between the corresponding brain states derived based on 236 ROIs and 360 ROIs.

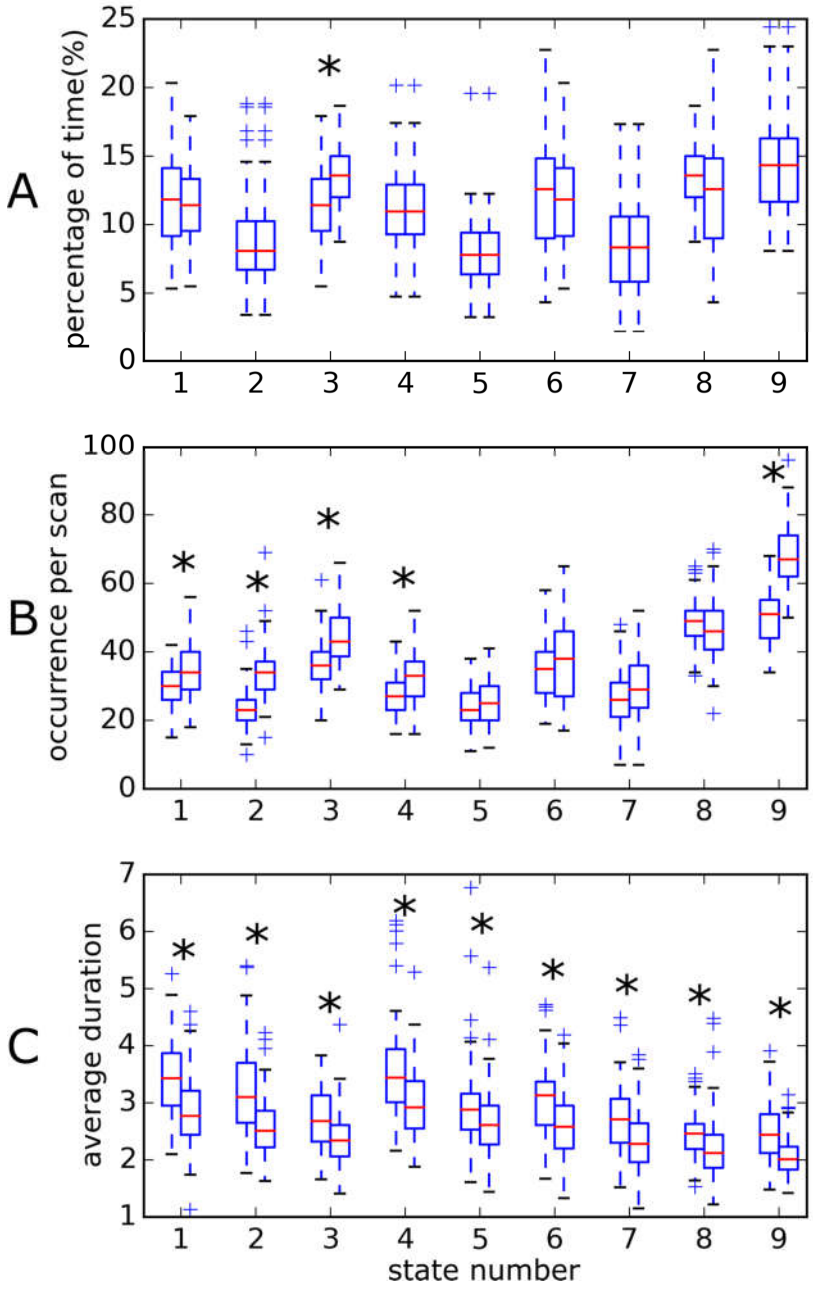

Figure 16 – The temporal characteristics between the corresponding brain states derived based on 236 ROIs (left) and 360 ROIs (right). The temporal characteristics are compared using boxplots of (A) time occupation, (B) occurrence, and (C) average duration. The star marks the significant difference between results from two atlases.

3.3.3 PAE

In order to identify the differences in brain state switch systems between healthy controls and individuals with PAE, GHMM was applied on both groups to derive 7 brain states and their characteristics according to the approaches described in the previous section. When two different GHMMs were trained on two different groups, we were not able to get the correspondence between the brain states across groups. We matched these states between two groups according to the similarity of spatial patterns. Note that, even if two states are paired, their correspondence is not guaranteed by the model. In fact, the pairing just represents their spatial similarity.

Figure 17 shows the spatial patterns of the brain states derived by two GHMMs. The first column is the 7 states of the control group, and the second column is the brain states of PAE group. Note that all the voxels shown are significantly activated or deactivated ($p < 0.05/total\ number\ of\ voxels$). S1 and S2 are whole brain deactivation and activation states in the control group, while for PAE, only the DMN is deactivated or activated. S4 and S5 in the control group are positive DMN plus negative TPN (+DMN-TPN) and +TPN-DMN, while in PAE group, this anti-correlation between DMN and TPN disappears and only the TPN shows deactivation and activation. The corresponding temporal characteristics are plotted side by side in Figure 18. There is no clear difference in the overall temporal characteristics.

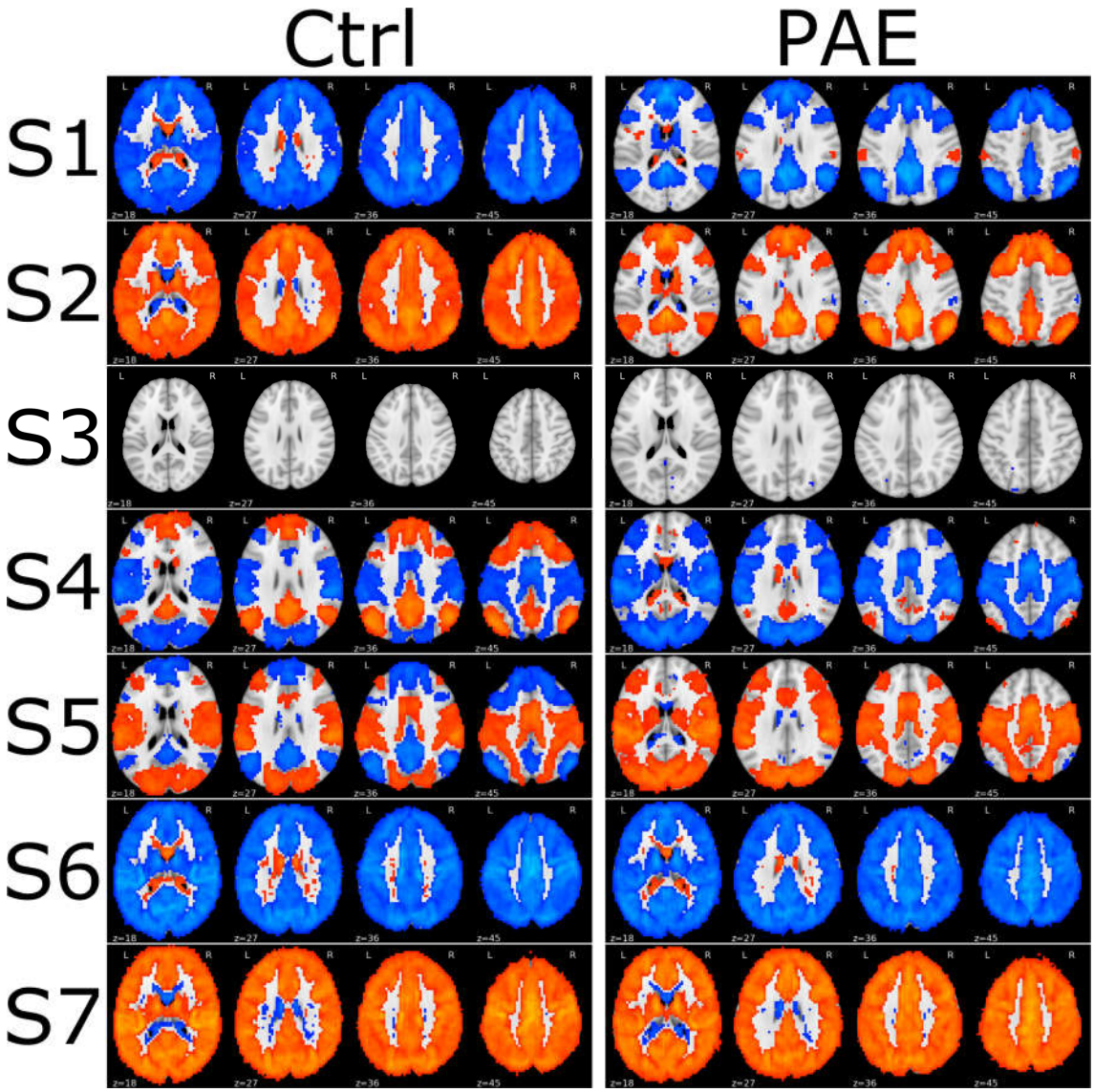

Figure 17 – The spatial patterns of brain states for control and PAE group. Note that only those voxels that are significantly activated or deactivated are color coded in the figure (p<0.05/total number of voxels).

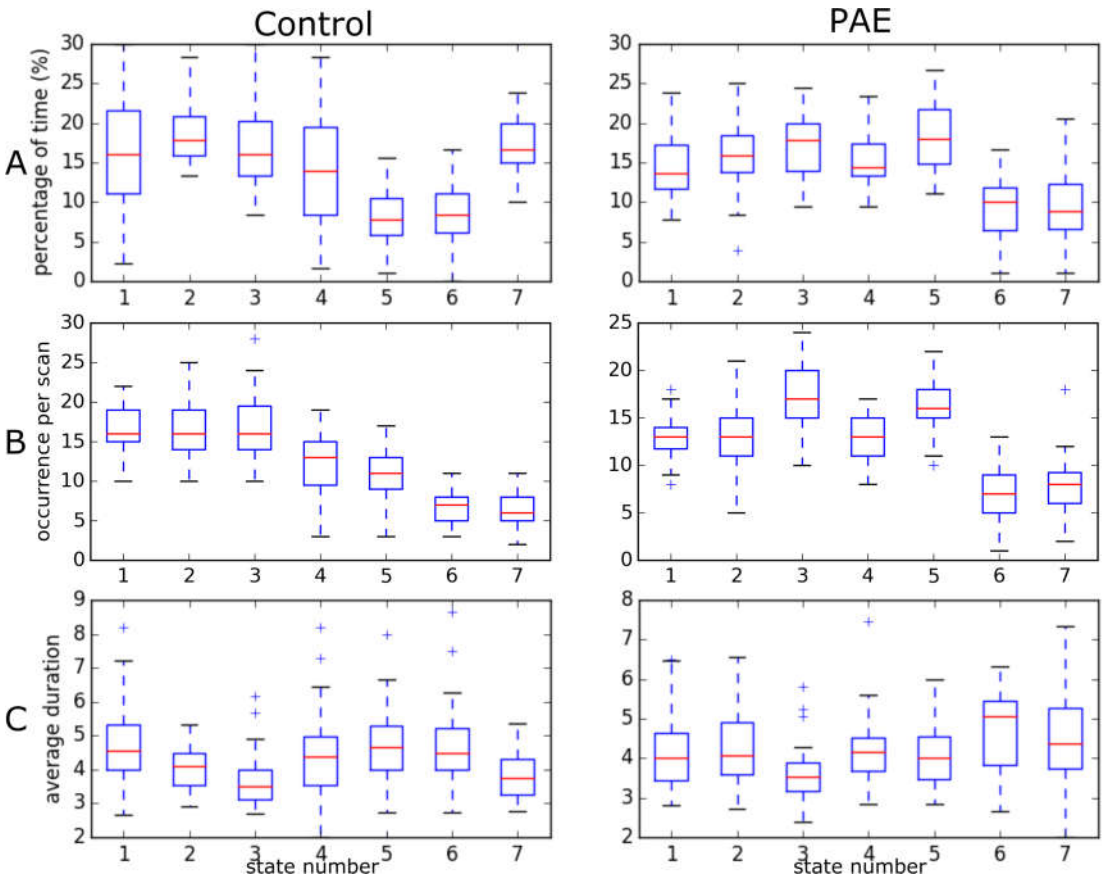

Figure 18 – The temporal characteristics between the corresponding brain states derived from control (left) and PAE (right) groups. The temporal characteristics are compared using boxplots of (A) time occupation, (B) occurrence, and (C) average duration.

Although we have shown some differences between the two groups, the correspondence of the brain state pairs is only based on the similarity of their spatial patterns. In the following results, we enforce the correspondence by only training one GHMM for all the subjects in both groups. The spatial patterns and temporal characteristics are generated based on the same GHMM but on different subject groups. Figure 19 and Figure 20 are the spatial patterns of brain states and the temporal features of brain states. Bonferroni correction was performed by dividing the total number of voxels or the total number of temporal features. A significant difference between states

only appears in some scattered voxels rather than large regions. Moreover, there is no significant difference of the temporal characteristics between groups. We have also tried a less restricted multiple-comparison-correction method, false discovery rate (Benjamini-Hochberg procedure (Benjamini and Hochberg 1995)), as shown in Figure 21. S4 shows a significant difference with a relatively large area in the dorsal lateral prefrontal cortex. Since we used 87 subjects to train the joint model, we were able to derive 9 brain states based on the data from two groups. The results are very similar to the results with 7 brain states as shown in Figure 22. The dorsal lateral prefrontal cortex shows the largest differences between groups. Note that when training with 9 brain states, some of the brain states became less stable.

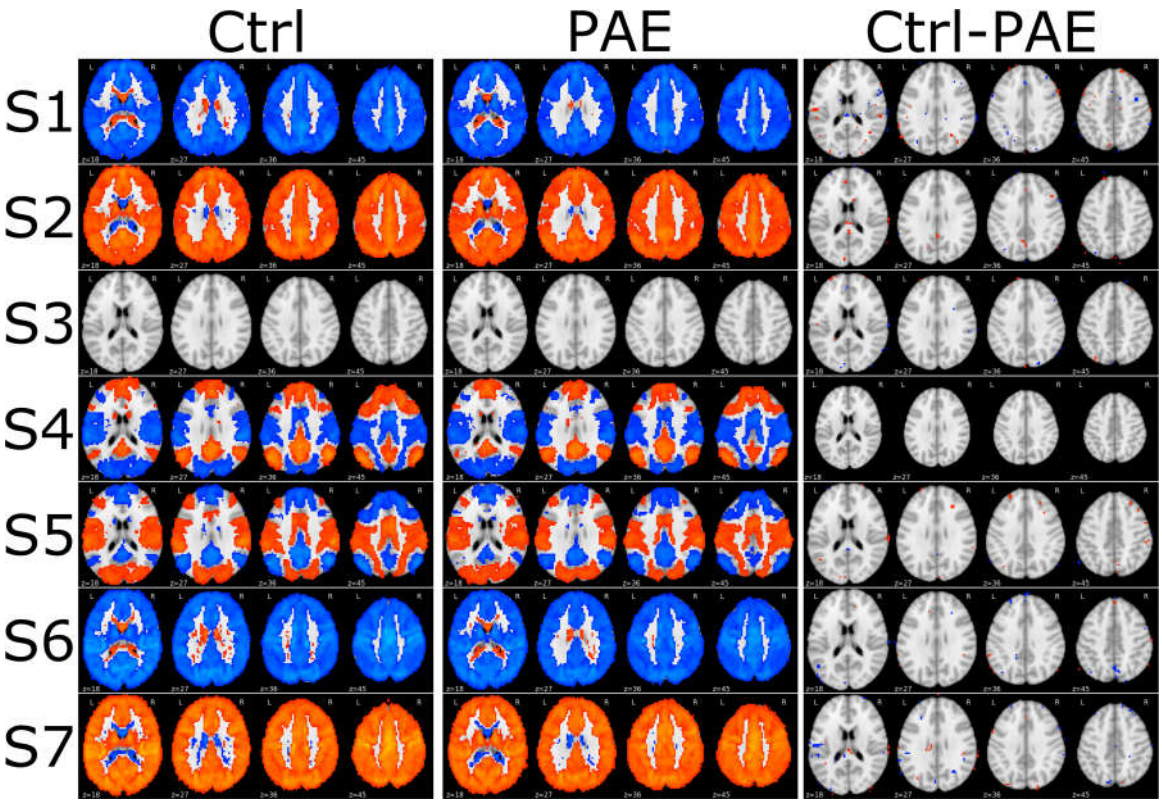

Figure 19 – The spatial patterns of brain states for the control and PAE groups. The difference between the groups is shown in the 3rd column. Note that only those voxels

that are significantly activated, deactivated, or differences between groups are color coded in the figure (p<0.05/total number of voxels)**.**

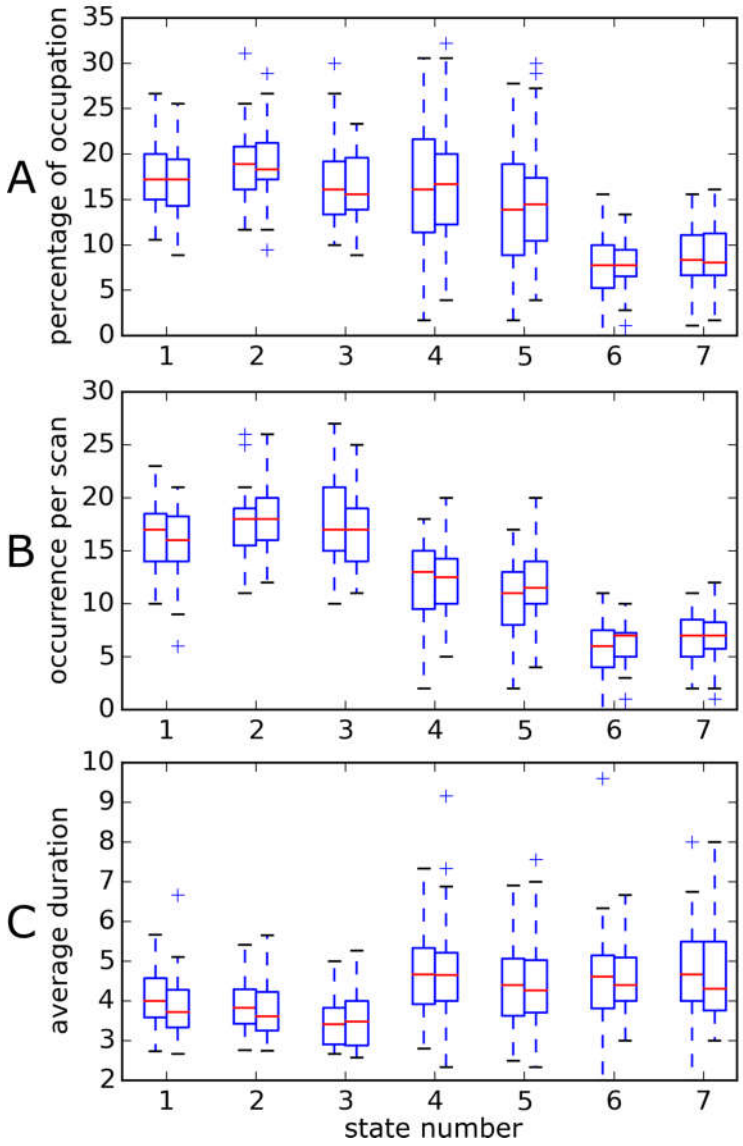

Figure 20 – The temporal characteristics between the corresponding brain states derived from control (left) and PAE (right) groups. The temporal characteristics are compared using boxplots of (A) time occupation, (B) occurrence, and (C) average duration. The stars indicate a significant difference between brain states (p<0.05/21).

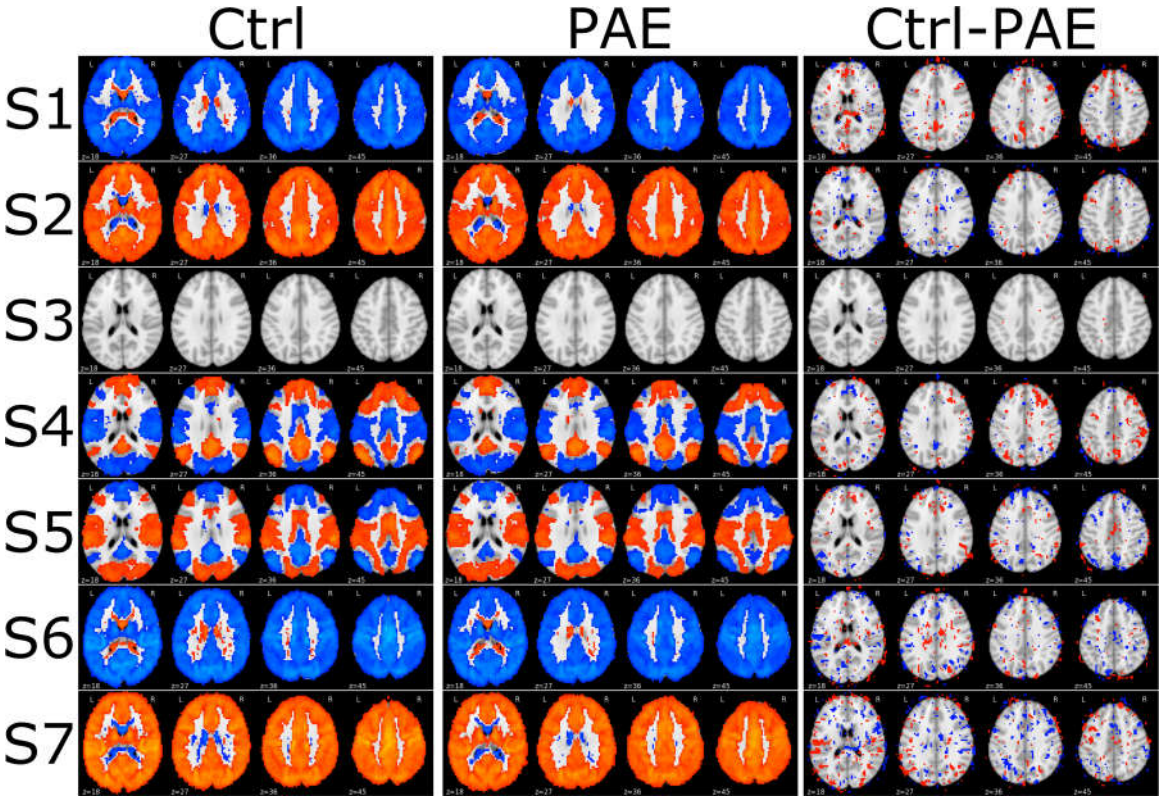

Figure 21 – The spatial patterns of brain states for control and PAE groups. The difference between the groups is shown in the 3rd column. Note that only those voxels that are significantly activated, deactivated, or differences between groups are color coded in the figure. The Benjamini-Hochberg procedure was adopted for multiple comparisons ($\alpha = 0.05$).

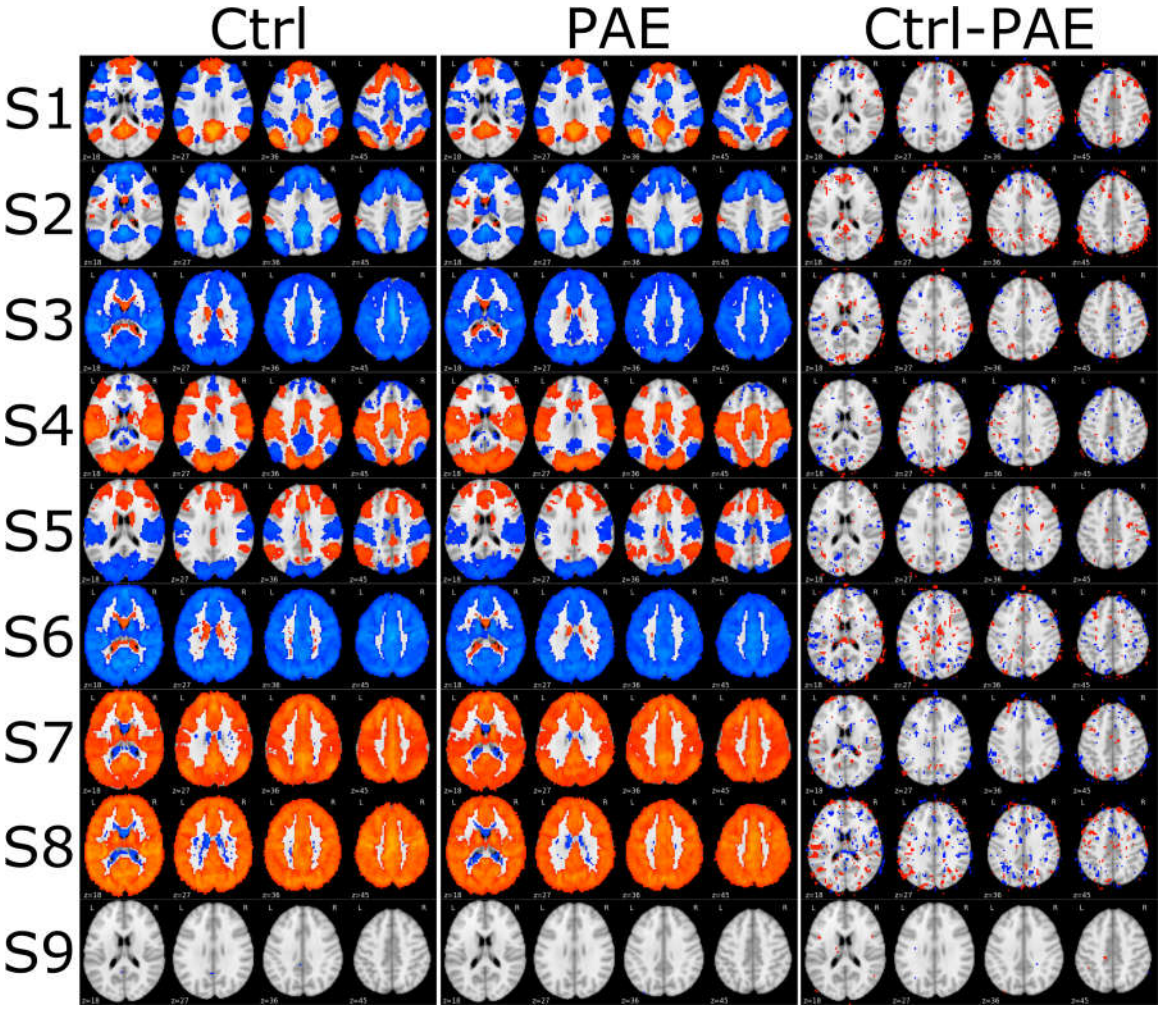

Figure 22 – The spatial patterns of brain states for control and PAE groups with 9 brain states. The difference between the groups is shown in the 3rd column. Note that only those voxels that are significantly activated, deactivated, or differences between groups are color coded in the figure. The Benjamini-Hochberg procedure was adopted for multiple comparisons ($\alpha = 0.05$).

3.4 Discussion

This chapter is an extension of the GHMM introduced in Chapter 2. We have further tested its reproducibility on different datasets, and the effects of different preprocessing steps including temporal filtering, GSR, and different atlases. We have also

applied GHMM on a PAE dataset to unveil the differences between the PAE and control groups. The implications of some results are discussed in the following paragraphs.

3.4.1 Reproducibility on different datasets

Our results indicate that GHMM is highly reproducible on different datasets, although the two datasets are greatly different from each other. The HCP dataset used a multi-band acquisition technique to achieve the state of the art spatial and temporal resolution: 2 mm × 2 mm × 2 mm and 0.72 s. The Beijing_Zang dataset used a very conservative scan protocol with no multiband acquisition. Therefore, the spatial and temporal resolution was lower: 3.1 mm × 3.1 mm × 3.6 mm and 2 s. The two datasets were also preprocessed differently. HCP data was processed using the minimal preprocessing pipeline of HCP, while Beijing_Zang data was processed with CPAC. Even the final data formats were different because the cortical area of HCP was surface based and Beijing_Zang was fully volumetric data. There were no significant differences in the subject groups, both were comprised of healthy young individuals. Some of the processing steps were applied to both datasets, which helped to reduce the difference in processing methodology. For example, they were both temporally filtered with the same bandpass frequency (0.01-0.1Hz) to reduce the difference caused by temporal resolution. Moreover, the time courses were extracted from atlases based on the same coordinates of nodes to decrease the spatial difference caused by spatial resolution and preprocessing.

Despite the difference of the datasets, the brain states are still very similar as shown in Figure 10. Because of the distinct data format, it is difficult to quantify the similarity of spatial patterns across datasets. However, most of the patterns look almost

identical. 8 out of 27 temporal features have significant difference across datasets (Figure 11). Overall, the spatial patterns are highly reproducible with different scan parameters and preprocessing approaches, while the temporal features are more affected by the difference in acquisition protocol and processing approaches.

3.4.2 Effect of temporal filtering

Because the hemodynamic response of neural activity is considered only in the low frequency range, a temporal bandpass filter (0.01-0.1Hz) is often applied to fMRI data to remove physiological or other non-neuronal information. Removing the high frequency components in fMRI will definitely affect the temporal features of the brain states as shown in Figure 13. Temporal filtering removes the high frequency signal and thus results in a smooth and slowly fluctuating signal. This explains why the state duration is significantly increased, while the occurrence of the states is significantly reduced ($p \ll 0.001$). This means the brain switches less often between states and stays in each state longer.

Temporal filtering can also lead to differences in the spatial patterns of brain states. We notice that in Figure 12, the "ground" state, which has no apparent activation or deactivation patterns, only appears after temporal filtering (S3). When the high frequency component in the signal is removed by temporal filtering, the transition between brain states is smoother and can be captured by the GHMM. When all the high frequency fluctuation is considered, this "ground" state might be too transient and is neglected by the model.

3.4.3 Effect of GSR

GSR can cause anti-correlation of multiple RSNs, an effect which has been noted since the earliest functional connectivity (Fox, Zhang et al. 2009), GSR caused anti-correlation of multiple RSNs in our study as well. For example, in the first two rows of Figure 14, S4 and S5 are mainly activation and deactivation in DMN before GSR; however, after GSR, they are positive DMN plus negative TPN and VN and their opposite. For S6 and S7, GSR also magnifies the anti-correlation between FPN and SMN. We also notice the whole brain activation and deactivation states, S8 and S9, disappear after GSR. These two patterns also resemble the global signal correlation maps reported in the literature (Power, Plitt et al. 2017) with SMN and VN being the most highly correlated regions. However, not all the anti-correlation is caused by GSR. We have already seen some anti-correlation of RSNs existing before GSR. For example, S1 and S2 show an anti-correlation between DMN and TPN, which is similar with and without GSR. Therefore, GSR eliminates the whole brain activation and deactivation states and magnifies the anti-correlation between some RSNs.

3.4.4 Impact of different atlases

Spatial patterns of brain states are not sensitive to differences in atlases, while the temporal features can be affected by the different number of ROIs. First, the two atlases we used are very different from each other as shown in Figure 23. The 236 ROIs are generated based on coordinates of functional nodes from a meta-analysis, and each ROI is 5mm in radius, which leaves some brain regions not included in any ROIs. The 360 ROIs, in contrast, are based on a parcellation of the brain, in which all the cortical regions are covered by one ROI. The spatial patterns of brain states derived based on these two atlases are almost identical despite the difference between two atlases (the second and

third row of Figure 14 as well as Figure 15). However, we notice that the occurrence and duration of almost all the brain states are affected by atlases. As we increase the number of ROIs, we also increase the dimension of the multivariate Gaussian distribution. This leads to more emphasis on the spatial pattern of brain states, which is represented by a Gaussian distribution, rather than the Markov switching process. As a result, the model tends to switch more often to match the spatial activity pattern, and this explains why the duration of all the states is significantly shorter and the occurrence is higher when using 360 ROIs.

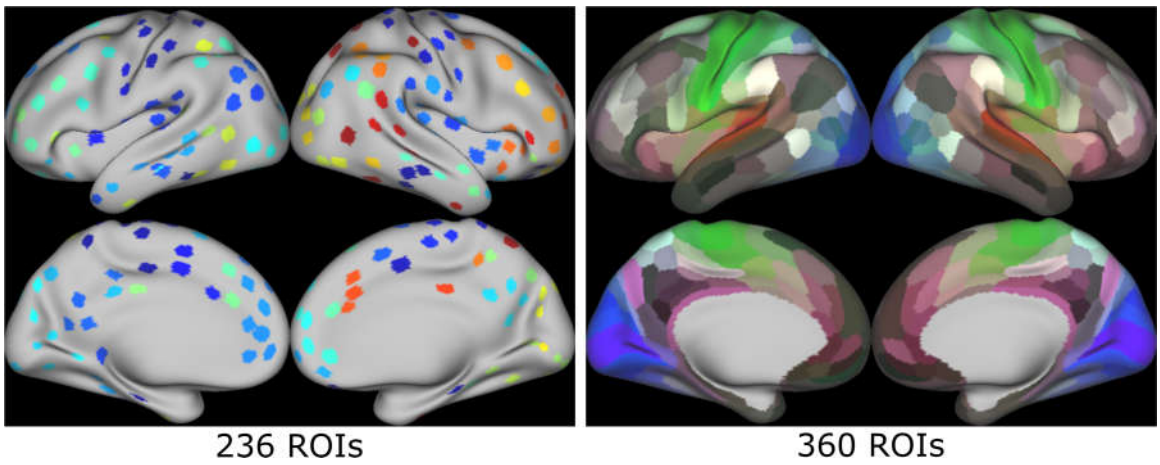

Figure 23 –Two atlases with 236 and 360 ROIs.

3.4.5 Effect of PAE

To increase the reproducibility of GHMM so that the differences between the PAE and control groups are truly the effect of PAE, the two groups of subjects were scanned with the same protocol, and the same preprocessing pipeline was applied on both groups to process the data.

The difference between groups was tested using two methods. We either trained two different GHMMs for the PAE and control groups, or one combined GHMM with all data from two groups. The results show that the differences in brain states caused by PAE can be detected on a system level (Figure 17 and Figure 18) rather than on a single state level (Figure 19 and Figure 20). We further notice that the relationship between DMN and TPN is weaker in the PAE group. For example, S4 and S5 in the control group shows anti-correlation between DMN and TPN; however, these two networks are separated into S1, S2, S4, and S5 in the PAE group. Given the high reproducibility of the GHMM, the difference between PAE GHMM and control GHMM suggests that PAE leads to brain function alteration. When training one GHMM on both the PAE and control groups, it is assumed that brain state switching systems are the same for the two groups. In this scenario, the difference on a single state level is not significant under Bonferroni correction (Figure 19 and Figure 20). When we switched to a less restricted multiple comparison correction method (the Benjamini-Hochberg procedure), a significant difference was present in the dorsal lateral prefrontal area, which is part of the DMN. Therefore, we conclude that PAE disturbs the whole brain switching system, especially the relation between DMN and TPN, and this effect is less detectable after training a combined model that includes a control group.

We have shown GHMM is capable of demonstrating differences between control and PAE subject groups. In future work, we plan to use GHMM as a classification tool to predict PAE in a given subject. In order to do that, we will adapt minimum classification error training to GHMM (Juang, Hou et al. 1997).

3.5 Conclusion

In this chapter, we have demonstrated the high reproducibility of GHMM by testing it on different datasets. We studied different preprocessing approaches and have shown three main effects. Firstly, temporal filtering influences both spatial and temporal characteristics of GHMM by removing the high frequency signal. Secondly, GSR removes the whole brain activation and deactivation states, and magnifies the anti-correlation between different RSNs. Finally, using different atlases does not change the spatial pattern of brain states; however, the temporal features are affected by the number of ROIs. We also applied GHMM to detect functional biomarkers of PAE to demonstrate that PAE disturbs the brain switching process on a system level; however, the effect is averaged out when the GHMM is trained on both the PAE and the control group.

CHAPTER 4. INDIVIDUAL IDENTIFICATION USING FUNCTIONAL BRAIN FINGERPRINT DETECTED BY RECURRENT NEURAL NETWORK

4.1 Introduction

In neuroimaging, researchers are fully cognizant of the uniqueness of individual's brain structure and have devoted a great deal of effort registering individual brains into a standard brain (Talairach and Tournoux 1988, Mazziotta, Toga et al. 1995). When performing group level analysis, functional signals of different subjects from the same brain region are often concatenated or averaged out, and the individual uniqueness is neglected. Individual differences in brain function have also been recently noted (Mohr and Nagel 2010). Moreover, functional connectivity (FC) has been used as a brain fingerprint which can identify unique characteristics of individuals (Finn, Shen et al. 2015). However, we believe that investigating the uniqueness in brain function could shed light on the underlying brain mechanism.

When considering FC as a personal trait, it is stable when enough data is acquired (600 time frames, 7.2 minutes). However, when only a short period of time is used (100 time frames, 72 seconds), the individual discriminating power is only about 70% on average (Finn, Shen et al. 2015). This could be due to the lack of statistical power when having fewer data and/or high variation of FC caused by the dynamics of FC (Allen, Damaraju et al. 2014). The dynamics of FC suggests that FC pattern derived based on a short time window can vary significantly during tens of seconds or several minutes

increasing the difficulties of individual identification. However, another possible explanation is that FC does not fully utilize the temporal information in the data since the temporal axis collapses when computing FC and only a spatial connectivity pattern is used to identify individuals. In order to test this hypothesis, we introduce a recurrent model that exploits both spatial and temporal information to learn and predict individual identification.

Recurrent neural network (RNN), a prevalent network structure in deep learning, is widely used in sequential learning, such as speech recognition and handwriting recognition (Mikolov, Karafiat et al. 2010, Graves 2012, Amodei, Anubhai et al. 2015). As data fed into RNN sequentially, the model will evolve over time based on its previous state and its current input. Gated recurrent unit (GRU) contains two gating units, i.e. update gate and reset gate, which enable the model to have a longer memory (Chung, Gulcehre et al. 2014). To the best of our knowledge, RNN and GRU have not hitherto been applied to identify individuals based on fMRI data.

In this study, we adapt a recurrent model based on GRU to investigate the individual uniqueness of brain function. We show that 100 time frames (72 seconds) of fMRI data provides sufficient enough information to identify individuals. Three different preprocessing approaches are used to study how global signal and different atlases affect the accuracy of individual identification. We also derive GRU patterns to visualize the important characteristics of brain function in terms of individual uniqueness.

4.2 Materials and methods

4.2.1 Dataset and preprocessing

FMRI data of 100 subjects from the HCP database (Van Essen, Smith et al. 2013) were used in the present work (age: 22-36^{+}, gender: 46M/54F, TR=0.72s). Each subject had 4 different scans with a total duration of 57.6 minutes (14.4 minutes×4). To reduce the anatomical differences, the preprocessed data, upon the application of the HCP preprocessing pipeline (Glasser, Sotiropoulos et al. 2013), was used in this work. More specifically, we used the data that was denoised with ICA-FIX (Griffanti, Salimi-Khorshidi et al. 2014, Salimi-Khorshidi, Douaud et al. 2014) and registered to group level cortical surface template with MSM-All (Robinson, Jbabdi et al. 2014). To decrease the computational complexity, signals from different regions of interest (ROIs) were extracted by averaging time courses of voxels inside the region. Two different atlases were used to test whether our approach is sensitive to atlases. For the first atlas, 236 ROIs with 5 mm radius over the cortices were generated based on a meta-analysis of fMRI (Power, Cohen et al. 2011). The second atlas has 360 ROIs derived based on multi-modal parcellation (Glasser, Coalson et al. 2016).

Data of the two scans conducted on the first day was adopted as the training dataset, and one scan from the second day was used as validation and the other scan of the second day was testing dataset. Each scan has 1200 time frames in total, and this entire scan was cropped into multiple 100-time-frame periods with a step size of 1 time frame. Within each 100-time-frame period, time series was demeaned and normalized within each 100-time-frame period before being fed to our recurrent learning model.

4.2.2 Recurrent learning model

The recurrent neural network was applied to capture the sequential information in fMRI data. Our deep learning model architecture has three layers as shown in Figure 24.

The first layer is an average pooling layer with a length of 4 on the temporal axis. This layer reduces computation complexity by shortening the number of time points from 100 time frames to 25 pooled time frames for the subsequent recurrent layer. Also, a 25% dropout was applied on the input to alleviate overfitting before the average pooling layer. The second layer is a recurrent layer with 256-dimension gated recurrent unit (GRU) as represented by the following equations. For simplicity, we will also refer the 256 dimension GRU as 256 GRUs:

$$
\begin{aligned}
z_t &= \sigma(W_z \cdot [h_{t-1}, x_t]) \\
r_t &= \sigma(W_r \cdot [h_{t-1}, x_t]) \\
\tilde{h}_t &= tanh(W_h \cdot [r_t h_{t-1}, x_t]) \\
h_t &= (1 - z_t) h_{t-1} + z_t \tilde{h}_t
\end{aligned}
\tag{3}
$$

where at time step t, x_t is the input of GRU, h_t is the output of GRU, z_t is the update gate and r_t is the reset gate. $\sigma(a) = 1/(1 + e^{-a})$ is the sigmoid function that maps the output value is between 0 and 1. Note that 25% dropout was also applied on both the input connections (from x_t) and recurrent connections (from h_{t-1}) to reduce overfitting. The recurrent layer will only generate output when it receives all the input data, which are 25 time frames from average pooling layer. This output is then connected to the top softmax layer with100 outputs to predict which individual the data is from.

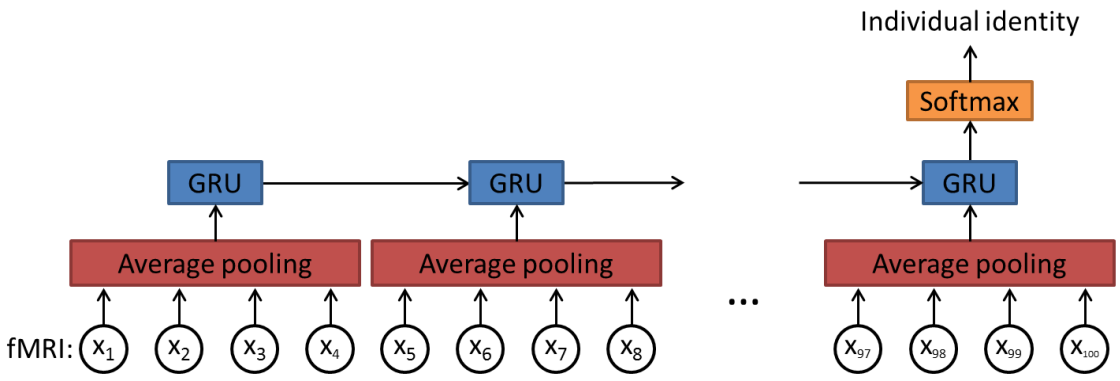

Figure 24 – The architecture of our three-layer recurrent neural network. A 25% dropout rate is applied to the input of average pooling layer as well as the input and recurrent connections of GRU layer.

Our recurrent learning model was built using Keras (Chollet 2015) and trained with Adam optimizer (Kingma and Ba 2014). Learning rate of the optimizer was set to 1×10^{-5}, and batch size was 64. The model was trained on the training dataset. At each epoch of training, the model was saved together with its loss calculated on validation datasets. The best model was the one with minimum validation loss. Finally, the performance of the best model was evaluated on the test dataset. Note that the test dataset was not involved during the training process.

4.2.3 Impact of preprocessing

There is evidence in the literature has shown that global signal of fMRI during resting state is associated with physiological signals, such as breath, heart rate, and head motion (Burgess, Kandala et al. 2016, Power, Plitt et al. 2017). In order to test whether our model is sensitive to these kinds of physiological information, we processed the data both with and without the global signal regression and compared the results. We also

tested whether our model was sensitive to atlas by extracting the time course from two different atlases with different number of ROIs as mentioned in the previous sub-section.

4.3 Results

When predicting individual identity using only 100 time frames (72 seconds) of fMRI data, our model was able to achieve over 90% accuracy on validation and testing data. In contrast, functional connectome fingerprint, which only uses FC, achieved less than 70% accuracy on average (Finn, Shen et al. 2015). Table 1 shows the accuracy of our model on three different preprocessing approaches (with and without GSR and from two different atlases). We performed a paired t-test between the accuracy with and without GSR to show GSR significantly increased test accuracy on data 236 ROIs (effect size = 0.51, $p = 1.37\times10^{-6}$). Paired t-test was also conducted between the accuracy with 236 and 360 ROIs to demonstrate that changing atlas affects the accuracy in a less significant way (effect size = 0.25, $p = 0.014$) on data with GSR.

Table 1 – The validation accuracy and test accuracy of three different processing approaches.

	Validation accuracy	Test accuracy
236 ROIs w/o GSR	92.80%	89.89%
236 ROIs w/ GSR	95.75%	94.43%
360 ROIs w/ GSR	94.14%	93.20%

The output of GRUs is shown in Figure 25. Dots with different colors represent data from different subjects. The output of GRU is a mapping from 256 dimensions to 2 dimensions by t-distributed stochastic neighbor embedding (t-SNE) (Maaten and Hinton 2008). As shown in the figure, data from different subjects are clustered together in the

second layer before going into the final classification layer, which is consistent with the high prediction accuracy.

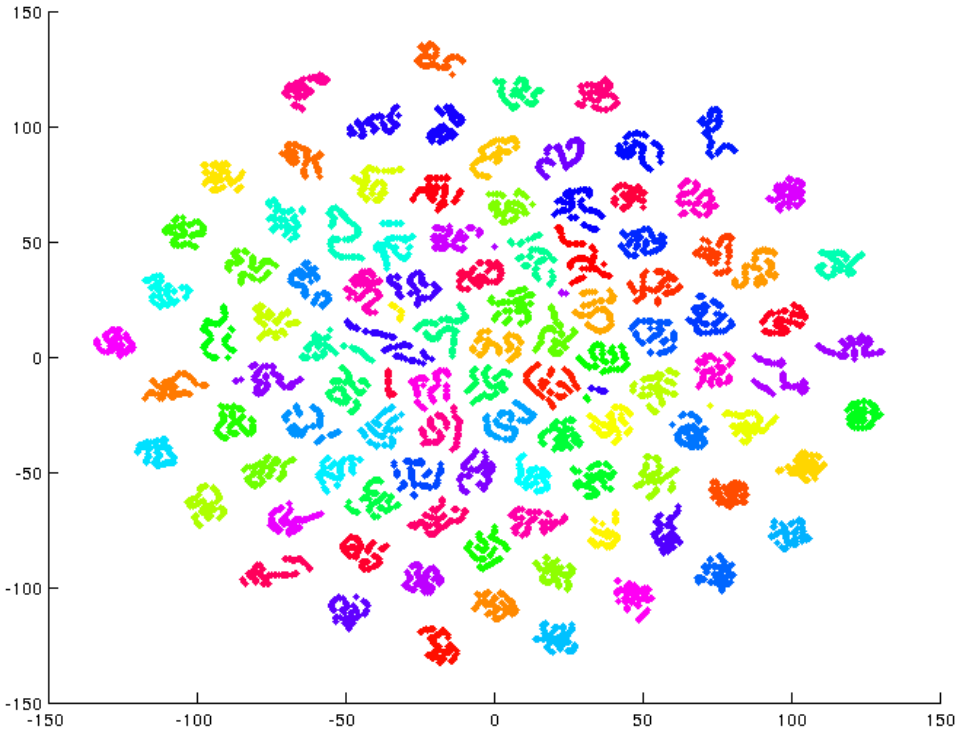

Figure 25 – Test data from different subjects are clustered as the output of the GRU. Dots with different colors are from different subjects. The 256-dimensional output of GRU is visualized in 2 dimensions by t-SNE.

Figure 26 shows the output of one GRU with respect to different subjects and time points. Note that the output is between -1 and 1 due to the $tanh$ function applied at the output of GRU. When a GRU is sensitive to 100 time frames of the input data, the GRU will output a high absolute value. Therefore, to discover what each GRU is sensitive to, we multiplied the rawdata with the corresponding output of GRU and summed the resultant spatial map of all the subjects over time. These weighted sum

spatial maps, referred to as GRU patterns, to characterize what is captured by GRUs and what is fed into to the classification layer as features.

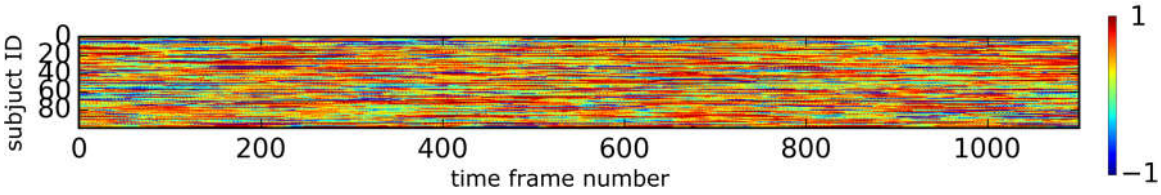

Figure 26 – The output of one GRU over time. Each row of the image represents the output of GRU with respect to a different subject. Each column of the image is the output of GRU at a particular time point.

The resultant 256 GRU patterns resemble RSNs in the literature (Yeo, Krienen et al. 2011, Lee, Smyser et al. 2013). Figure 27 shows 5 representative GRU patterns corresponding to task positive network (TPN), default mode network (DMN), frontoparietal network (FPN), visual network (VN), and somatosensory motor network (SMN). We also noticed that variations of one RSN could be captured by multiple GRUs. In order to cluster similar GRU patterns together and understand the features captured by GRUs in general, we used hierarchical clustering with cosine distance metric and used 0.8 as the threshold for demonstration purpose as shown in Figure 28. The threshold was chosen such that we had some diversity in the resultant clusters and not too many clusters with similar patterns. The cluster centers were computed by averaging all the GRU patterns in each cluster.

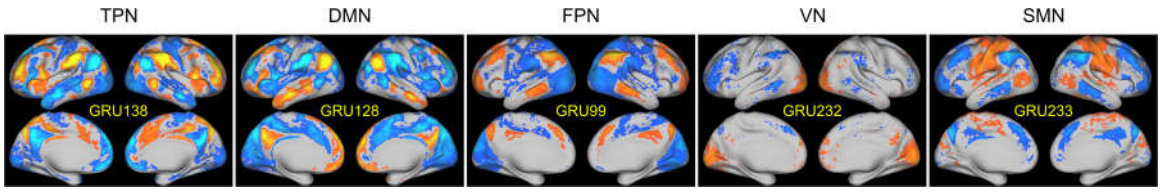

Figure 27 – Five representative GRU patterns that resemble 5 different RSNs.

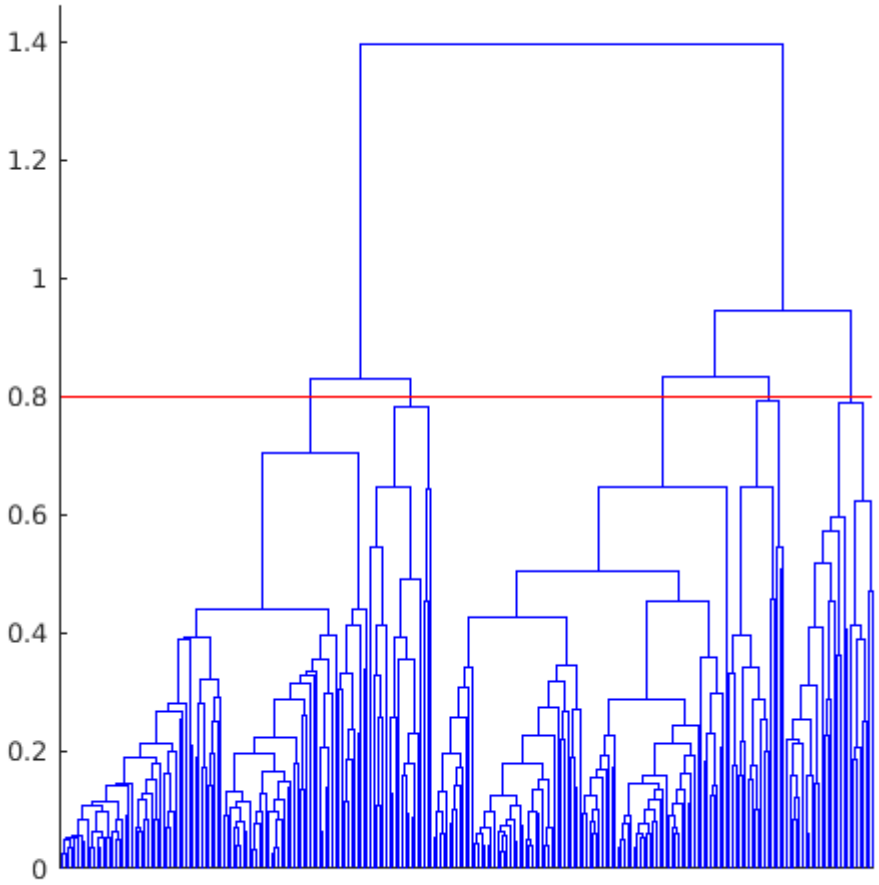

Figure 28 – Dendrogram of hierarchical clustering to group 256 GRU patterns into 5 clusters. The red line indicates the threshold to separate all GRU patterns into clusters.

The same approach was applied on the 3 differently preprocessed datasets to derive cluster centers of GRU patterns. This resulted in 5, 10, and 8 cluster centers on datasets of 236 ROIs without GSR, 236 ROIs with GSR, and 360 ROIs with GSR as shown in Figure 29. Some of the clusters only have a few GRU patterns in them; therefore, the average patterns of those clusters appear to be less smooth compared to the

rest. If we only focus on what most of the GRUs are sensitive to, Figure 30 summarizes those clusters that have more than 10 GRU patterns.

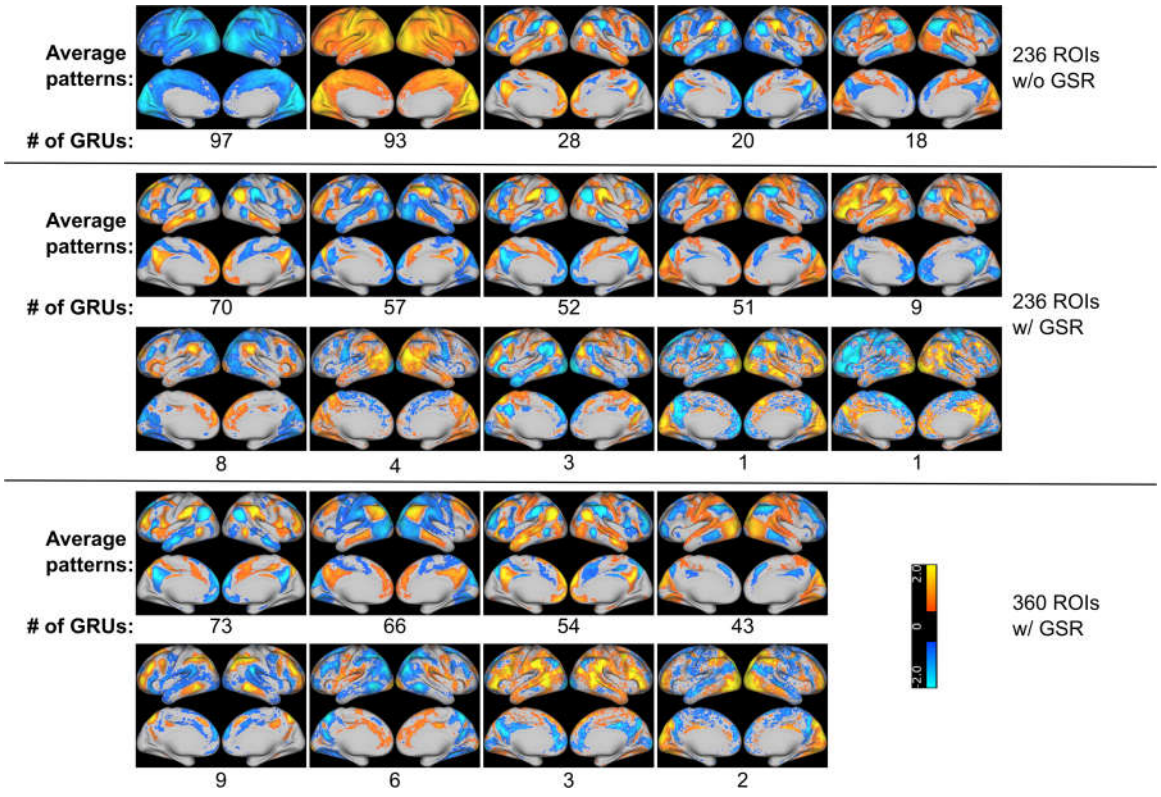

Figure 29 – Cluster centers of GRU patterns with 3 different preprocessing approaches. The number below the image indicates the number of GRU patterns assigned to each cluster.

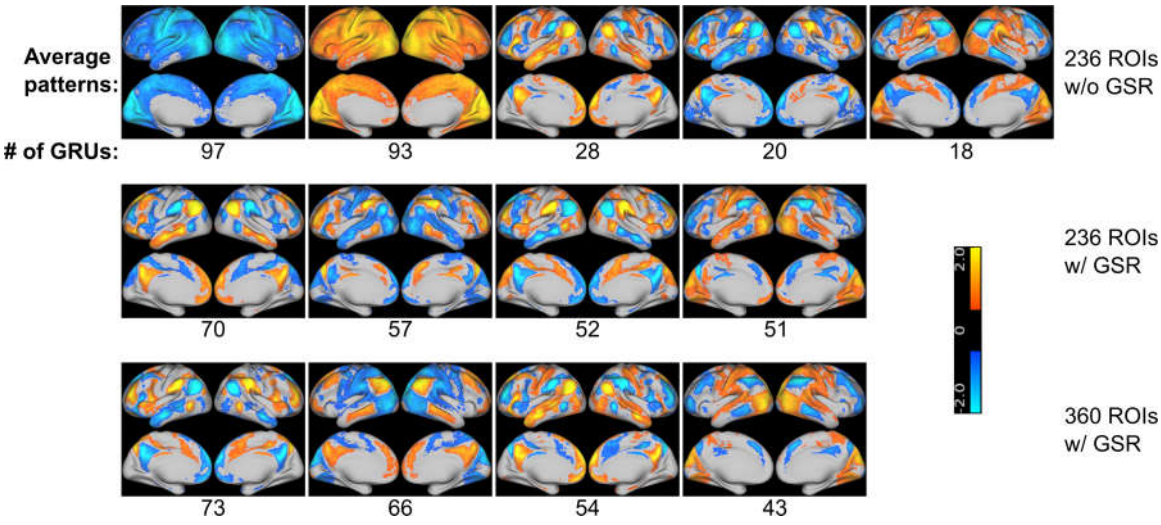

Figure 30 – Cluster centers of GRU patterns with 3 different preprocessing approaches. The number below the image indicates the number of GRU patterns assigned to each cluster. Only clusters with more than 10 GRU patterns are shown here.

We noted that when and only when GSR was not performed, most of the GRUs captured whole brain deactivation and activation as shown by the first two images in the first row of Figure 30. In these two whole brain deactivation/activation images, somatosensory, motor, visual cortices appear to be the most deactivated/activated. Among all three different preprocessing approaches, positive DMN plus negative TPN (+DMN-TPN) and positive TPN plus negative DMN (+TPN-DMN) are captured by most GRUs. The number of GRUs capturing these two patterns is listed in Table 2. Also, 18 GRUs detect positive SMN and VN together with negative DMN when GSR was not used. In contrast, 51 and 43 GRUs detect positive SMN and VN plus negative TPN after GSR with 236 and 360 ROIs respectively. Negative SMN and VN are also captured by GRU after performing GSR. 57 GRU patterns show negative SMN and VN plus positive TPN when using 236 ROIs, while 66 GRU patterns appear to be negative SMN and VN plus positive DMN when using 360 ROIs.

Table 2 – The number of GRUs that capture +DMN-TPN and +TPN-DMN patterns under 3 different preprocessing approaches.

	+DMN-TPN	+TPN-DMN
236 ROIs w/o GSR	28	20
236 ROIs w/ GSR	70	52
360 ROIs w/ GSR	54	73

We have also applied our model in gender prediction with HCP subjects. A total number of 803 subjects were used (600 for training, 100 for validation, 103 for testing). Following the same procedure, we were able to achieve a validation accuracy of 81.38%

and a test accuracy of 74.34%. The top five important features that were used to predict gender are shown in Figure 31. Among all the RSNs, SMN and FPN are the most valuable features for gender prediction.

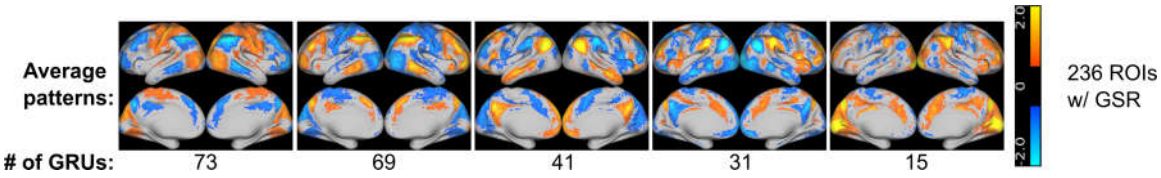

Figure 31 – Cluster centers of GRU patterns for gender prediction. The number below the image indicates the number of GRU patterns assigned to each cluster. Only clusters with more than 10 GRU patterns are shown here.

4.4 Discussion

A lot of effort has been devoted to reducing structural differences of individuals, such as developing brain template, volumetric registration, and surface registration (Robinson, Jbabdi et al. 2014) so that group level analysis can be conducted. However, we have shown that individual variability exists in how the brain functions over time. Given only 72 seconds of fMRI data, our recurrent model was able to predict the identity of each individual with 94^{+}% accuracy. The implications of our results are discussed in the following paragraphs.

We have shown that 100 time frames provide sufficient information to identify individuals, and our recurrent model outperforms the FC based approach in individual identification with 72s of fMRI data (94% vs 70%) (Finn, Shen et al. 2015). The FC based approach can be considered as handcraft feature selection plus k-nearest neighbor classifier ($k = 1$), while our recurrent model takes time series of multiple ROIs as input and learn its own features, i.e. GRU patterns, to make prediction. The dynamic nature of

FC can be one explanation of poor performance of FC based approach especially when only using 72s of data. However, it is also possible that the handcraft features, FC patterns, do not fully utilize the temporal information of the data, because the temporal axis is collapsed in FC and only the spatial information is preserved. In contrast, our recurrent model is able to use both spatial and temporal information of the data to derive better features to achieve the identification task.

4.4.1 Duplications of GRU patterns

The dropout layers in our recurrent model are the cause of the duplication in GRU patterns, while it also helps us understand which features are important. Dropout layer was introduced to deep learning to reduce overfitting by randomly setting some of the input dimension to zeros while preserving the rest (Srivastava, Hinton et al. 2014). In our model, dropout was applied on input layers as well as input and recurrent connections. If the feature is important, rather than entirely rely on one unit, the model will create multiple copies of the important features with some variations in GRU patterns just in case some GRUs are randomly set to zeros. This explains why many of our GRU patterns are similar. Actually, the number of GRU patterns that show a similar RSN reflects how important that RSN is when deciding the individual identity, i.e. how much the final decision relies on a particular RSN.

4.4.2 Effect of anatomical differences

Although data of all subjects are registered to a template, completely removing the anatomical difference is still difficult. To reduce anatomical contribution to individual difference, we used the data after volumetric and surface-based registration. The surface

registration applied on the HCP data, MSM-ALL, is a multimodal surface matching algorithm which also considered functional scans. Table 3 shows a great decrease in accuracy of the model when switching from resting state to language task fMRI. This decrease of accuracy is an evident suggesting that the individual discriminating power is based on brain function rather than purely anatomical differences of the subjects because the brain structure of the subject is the same between resting state and task. In fact, there is even no reposition of the subject between these two scans, because the language task and REST2 were both acquired on the same day. Several other aspects related to anatomical difference, such as partial volume effect and choice of atlases, are also worth noting.

We believe partial volume effect does not benefit individual identification. Since training data are from the first day and validation and testing data are from the second day, the repositioning of the subject could have created a different partial volume effect. Moreover, the ROIs we used were 5 mm in radius and would have lessened the partial volume effect. Therefore, the individual identification ability is unlikely arising from the partial volume effect.

We also tested whether the model is sensitive to the choice of atlases by using two sets of completely different ROIs. The 236 ROIs are generated based on coordinates of functional nodes from a meta-analysis, and each ROI has 5 mm radius, which leaves some of the brain region not covered by any ROIs. The 360 ROIs, in contrast, is based on a parcellation of the brain, in which all the cortical regions are accounted for. The 2nd and 3rd rows of Table 1 indicate that the difference in atlases did not substantially affect the accuracy. Although the difference is statistically significant (p = 0.015), its effect size is

much lower compared to the difference caused by GSR (0.25 vs. 0.51). The 2nd and 3rd rows of Figure 30 show that the features extracted by GRUs are also very similar between the two atlases. They both have +DMN-TPN, +TPN-DMN, and +SMN+VN-TPN patterns. When using 236 ROIs, GRUs tend to capture +TPN-SMN-VN, while with 360 ROIs, GRUs capture +DMN-SMN-VN. Therefore, the change of atlases does affect the GRU patterns a little; however, most of the RSNs captured by GRUs remain the same, and this change does not affect the accuracy greatly.

Table 3 – The test accuracy on resting state and language task fMRI of three different processing approaches.

	Test accuracy on REST2	Test accuracy on LG
236 ROIs w/o GSR	89.89%	55.09 %
236 ROIs w/ GSR	94.43%	66.29%
360 ROIs w/ GSR	93.20%	60.05%

4.4.3 Effect of physiological signal

Physiological information, such as respiration, heartbeat, and head motion affects the fMRI signal (Burgess, Kandala et al. 2016). For example, the effect of head motion was shown to be consistent across different scans of the same subject and correlated with some behavioral measurements (Siegel, Mitra et al. 2016). Meanwhile, this physiological signal has also been shown to be correlated with the global signal of fMRI (Power, Plitt et al. 2017). Therefore, to test whether our model is based on neuronal or physiological information, we compare the results before and after GSR. If our model substantially relies on this physiological signal, we expect to see a performance decrease after conducting GSR. Although GSR is not guaranteed to eliminate all the physiology related

variations, the fact that reducing this physiological signal actually increase the prediction accuracy suggests that our model does not heavily rely on physiological information to predict individual identification. In addition, most of the GRU patterns we derived resemble the conventional RSNs as shown in Figure 30. This also implies that our model is based on neuronal rather than physiological information to identify individuals. When comparing the GRU patterns derived without GSR and with GSR, the two global activation and deactivation patterns (first two images in the first row of Figure 30) are similar to the global signal correlation maps reported in literature (Power, Plitt et al. 2017) where somatosensory, motor, visual cortices are shown to be most correlated with global signal. This physiological fluctuation in global signal affects the features captured by GRU, and thus diminishes the prediction power.

4.4.4 Limitations and future directions

Several limitations of the methodology and additional considerations should be noted. First, our results suggest that anatomical information was not the main source for individual identification power; however, volumetric and surface-based registration is not guaranteed to eliminate all anatomical differences. How much of the discriminating power arises from anatomical information needs further investigation. Second, GSR cannot assure the complete removal of the physiological signal. In fact, head motion is shown to have a local effect on fMRI signal (Burgess, Kandala et al. 2016, Siegel, Mitra et al. 2016). It needs to be further examined how this local effect of the physiological signal can affect the results. For future directions, we believe our model can also be used in behavioral prediction or classification of brain diseases. For example, we can predict the individual traits, such as gender (as shown in Figure 31), age, and fluid intelligence,

and visualize and sort the features get captured by the model. In this way, we can gain a better understanding of the relationship between these individual traits and brain function. With enough data, we can start investigate how brain affects the brain by performing disease classification and study the features captured by the model.

4.5 Conclusion

We have introduced recurrent neural network to individual identification based on resting state fMRI data. We claim that 100 time frames (72s) are sufficient for training and predict individual identity with over 90% accuracy. By analyzing data with 3 different preprocessing approaches, we were able to show that our recurrent model can utilize neuronal information in fMRI data to extract GRU patterns and identify subjects. Therefore, we conclude that the brain functional fingerprints, i.e., spatiotemporal features of brain function, can improve our understanding of individual brains' function.

CHAPTER 5. CONCLUSION AND FUTURE WORK

We have introduced two machine learning models, GHMM and RNN, to systematically model brain temporal dynamics. For each of the two models, corresponding visualization and quantification of brain states has also been performed. With these approaches, we can compare differences in brain states caused by diseases, as well as individual and gender differences, visually and quantitatively. We have also broached several controversial preprocessing topics, such as global signal regression and temporal filtering. We have shown how these different steps affect the spatial and temporal patterns of brain states, and how they influence individual identification accuracy.

5.1 Differences between GHMM and RNN

Two models adopted in this work, GHMM and RNN, are both spatiotemporal models based on brain activation sequences; however, there are differences in the implementation of the models. For the methodology, GHMM has a finite number of hidden states which need to be specified before training. RNN, however, uses a high dimensional vector to represent its hidden states, which in theory can have an infinite number of hidden states, although the dimension of the hidden state vector is pre-defined. GHMM assumes the brain switches between states as a stochastic process, while the hidden state evolution of RNN is a deterministic process. For GHMM, the current brain state is only dependent on the previous state, which means the system does not have long term memory. For RNN, the current brain state depends on its previous state and current input as well, but with the update gate and reset gate, GRU theoretically can carry

memory indefinitely. The training processes of these two models are also different. We trained GHMM in an unsupervised learning fashion, which means we tried to find the model parameters that best fit the fMRI data. On the other hand, RNN, like many other deep learning models, is trained in a supervised learning fashion. This means that during training, both the fMRI data and the label (subject identity or gender) are fed into the model together so that it can find the best parameters to predict the individual label. Note that, since our RNN model has many more parameters than our GHMM, it requires a much larger dataset to achieve a satisfactory prediction performance.

5.2 Similarities between GHMM and RNN

Despite the differences in methodology, the brain states we derived from the two models share some commonalities. Note that in Chapter 4, we referred to the brain states derived by RNN as "GRU patterns" to emphasize that they are features learned by GRUs. First, the number of brain states in both models is similar. In GHMM, we restricted the total number of states to be 9 in order to achieve 100% algorithm stability, while in RNN, we derived 256 GRU patterns at first, then grouped them into 5, 8, and 10 brain states based on their similarity. This small total number of states is similar to that used in many other studies (Allen, Damaraju et al. 2014, Taghia, Ryali et al. 2017). Interestingly, we also see that small variations of the same brain state can also be captured by RNN. For example, in Figure 29 (2^{nd} row, 1^{st} column), we have 70 GRU patterns that look like activated DMN plus negative TPN. These variations in brain states are shown to be important when identifying individuals, which indicates that these variations could be a result of individual differences. It would be interesting to account for the contribution of individual differences and investigate whether some of these variations should be further

separated into different brain states. Second, most brain state spatial patterns derived by the two models are almost identical. For example, positive DMN plus negative TPN and negative DMN plus positive TPN are detected by both GHMM and RNN (S6 and S7 in Figure 3, 1st and 3rd column of 2nd row in Figure 29). This pair of brain states has been reported in other activation based studies (Liu, Chang et al. 2013). Also, activation and deactivation states over the whole brain are captured by both models (S1 and S2 in Figure 3, 2nd and 1st column of 1st row in Figure 29) before performing global signal regression. The fact that these two methodologically different models derive very similar brain states indicates that the detected brain states play an essential role in characterizing brain function.

5.3 Thesis contribution

One contribution of this work is that by systematically modeling the brain state switching process, we can link the derived brain states with brain diseases and individual traits. For example, by training two GHMMs on PAE and control groups, we were able to show how PAE has altered the brain state switching process by reducing the anti-correlation between the DMN and TPN networks. With enough training data, these two resulting models could be used as a generative model to examine which model fits a subject's data better and thus classify whether a subject is PAE or not. There also exists advance training methods to further improve classification accuracy (Juang, Hou et al. 1997) using the same model. As we mentioned in the introduction section of Chapter 3, the brain function alterations caused by PAE detected by stationary approaches are among multiple RSNs, including DMN, LPFN, RPFN, PMN, PVN, and ESVN. Some of these results contradict each other. Our model, however, inspects this problem from a

dynamic perspective, and most of the differences in brain states can be explained by the decoupling of DMN and TPN. It would be interesting to see whether other dynamic approaches will have similar results to ours. As another example, by identifying subjects and their gender, our RNN is able to associate different RSNs with different individual traits (Identity: DMN and TPN; gender: SMN). Non-neuronal signal, such as head motion and physiological fluctuation is known to contaminate the BOLD signal and is highly correlated with global signal. Our results in Table 1 suggest that these non-neuronal signals do not contribute greatly to inter-subject differences, while neuronal signal is what establishes an individual's uniqueness. However, global signal can still potentially reveal group level differences as it is shown to be related with vigilance and arousal level (Wong, Olafsson et al. 2013).

Our work has also provided additional information about which metrics characterize brain states better. Brain states can be defined as spatial activation or connectivity patterns that remain relatively stable for a short period of time (a few seconds to minutes). From the definition, there exists two different metrics that can represent a brain state. These are the spatial activation pattern of the brain (Liu, Chang et al. 2013, Liu and Duyn 2013), and connectivity between different brain regions (Allen, Damaraju et al. 2014). Both models introduced in this work utilize activation pattern of fMRI data to derive brain states. In Chapter 4, we demonstrated that our activation based model is better than a functional connectivity based approach in terms of individual identification accuracy (Finn, Shen et al. 2015). This result suggests that activation pattern based brain states can reveal more information than functional connectivity based brain states. Moreover, functional connectivity based analysis usually needs a pre-defined

fixed window length to compute correlation matrices. The length of the window should approximate the length of brain states which is often difficult to estimate (Shakil, Lee et al. 2016) beforehand. As shown in Chapter 2, our activation based model is able to detect brain states without a pre-defined window. Furthermore, we can compute the duration of each brain state after decoding the brain state transition sequence (Figure 4). Our results suggest that different brain states may have different durations. Also, each time a brain state occurs, it may have a slightly different duration. These findings all indicate that the fixed window length in functional connectivity based analysis can mix signals from two different brain states into one window, leading to signal cancellation. Therefore, an activation pattern based approach, which does not need a predefined window, offers more freedom in brain state characterization.

5.4 Limitations and future work

A few aspects regarding brain state metrics are worth investigating in the future. First, one functional connectivity pattern usually corresponds with two activation patterns. For example, the anti-correlation of DMN and TPN in functional connectivity will match with two opposite activation patterns: positive DMN plus negative TPN and negative DMN plus positive TPN (S6 and S7 in Figure 3). Moreover, the duration of these functional connectivity based brain states is on the order of minutes, while the duration of our activation based brains state is on the order of seconds. Therefore, it would be interesting to see whether the brain often switches between opposite brain activation patterns. This switching between opposite activation pattern may also involve some intermediate states, such as S9 in Figure 3. Further investigating and comparing the state transition sequence of both functional connectivity and activation based brain states

may explain the difference in brain state duration of the two different approaches and improve our understanding of how the brain functions over time. In Chapters 2 and 3, we used a multivariate Gaussian distribution to model the brain states, where the covariance matrix is restricted diagonally. This forces our model to capture activation based brain states. It has been shown that without the diagonal constraint, GHMM is able to capture functional connectivity based brain states (Eavani, Satterthwaite et al. 2013, Taghia, Ryali et al. 2017). The state sequence of this model will indicate what the duration of functional connectivity based brain states is, and how much it varies over time. It would be interesting to see what brain states can capture if we use a more complex emission distribution, such as a Gaussian mixture model. For example, several DMN activation patterns with some variations were presented in CAPs (Liu, Chang et al. 2013, Liu and Duyn 2013). With a Gaussian mixture model, these variations of the same RSN can be grouped into one single brain state and presented in a hierarchical way.

In addition to the limitations of each model which is mentioned in each chapter, one limitation for current fMRI analysis is that BOLD signal is an indirect measurement of neuronal activity. This limitation hinders the interpretation of solely fMRI-based results. We believe future multimodal studies can overcome this limitation and deepen our understanding of the BOLD signal, which will further improve fMRI preprocessing by filtering out more nuisance signal.

A few large cohort studies of brain diseases have already been started, such as UK biobank and Rhineland. Each of these studies will recruit tens of thousands to hundreds of thousands subjects. This huge amount of data leveraged by deep learning models will improve our understanding of brain diseases.

REFERENCES

Allen, E. A., E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele and V. D. Calhoun (2014). "Tracking whole-brain connectivity dynamics in the resting state." Cereb Cortex **24**(3): 663-676.

Amodei, D., R. Anubhai, E. Battenberg, C. Case, J. Casper, B. Catanzaro, J. Chen, M. Chrzanowski, A. Coates and G. Diamos (2015). "Deep speech 2: End-to-end speech recognition in english and mandarin." arXiv preprint arXiv:1512.02595.

Avants, B. B., N. J. Tustison, G. Song, P. A. Cook, A. Klein and J. C. Gee (2011). "A reproducible evaluation of ANTs similarity metric performance in brain image registration." Neuroimage **54**(3): 2033-2044.

Baker, A. P., M. J. Brookes, I. A. Rezek, S. M. Smith, T. Behrens, P. J. P. Smith and M. Woolrich (2014). "Fast transient networks in spontaneous human brain activity." Elife **3**: e01867.

Beckmann, C. F., M. DeLuca, J. T. Devlin and S. M. Smith (2005). "Investigations into resting-state connectivity using independent component analysis." Philosophical Transactions of the Royal Society of London B: Biological Sciences **360**(1457): 1001-1013.

Behzadi, Y., K. Restom, J. Liau and T. T. Liu (2007). "A component based noise correction method (CompCor) for BOLD and perfusion based fMRI." Neuroimage **37**(1): 90-101.

Benjamini, Y. and Y. Hochberg (1995). "Controlling the false discovery rate: a practical and powerful approach to multiple testing." Journal of the royal statistical society. Series B (Methodological): 289-300.

Bilmes, J. A. (1998). "A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models." International Computer Science Institute **4**(510): 126.

Biswal, B. B., M. Mennes, X.-N. Zuo, S. Gohel, C. Kelly, S. M. Smith, C. F. Beckmann, J. S. Adelstein, R. L. Buckner and S. Colcombe (2010). "Toward discovery science of human brain function." Proceedings of the National Academy of Sciences **107**(10): 4734-4739.

Broyd, S. J., C. Demanuele, S. Debener, S. K. Helps, C. J. James and E. J. Sonuga-Barke (2009). "Default-mode brain dysfunction in mental disorders: a systematic review." Neuroscience & biobehavioral reviews **33**(3): 279-296.

Buckner, R. L., J. R. Andrews-Hanna and D. L. Schacter (2008). "The brain's default network." Annals of the New York Academy of Sciences **1124**(1): 1-38.

Burgess, G. C., S. Kandala, D. Nolan, T. O. Laumann, J. D. Power, B. Adeyemo, M. P. Harms, S. E. Petersen and D. M. Barch (2016). "Evaluation of denoising strategies to address motion-correlated artifacts in resting-state functional magnetic resonance imaging data from the Human Connectome Project." Brain Connectivity **6**(9): 669-680.

Chollet, F. (2015). Keras. Available from: https://github.com/fchollet/keras.

Chung, J., C. Gulcehre, K. Cho and Y. Bengio (2014). "Empirical evaluation of gated recurrent neural networks on sequence modeling." arXiv preprint arXiv:1412.3555.

Coles, C. D. and Z. Li (2011). "Functional neuroimaging in the examination of effects of prenatal alcohol exposure." Neuropsychology review **21**(2): 119-132.

Cox, R. W. (1996). "AFNI: software for analysis and visualization of functional magnetic resonance neuroimages." Computers and Biomedical research **29**(3): 162-173.

Eavani, H., T. D. Satterthwaite, R. E. Gur, R. C. Gur and C. Davatzikos (2013). Unsupervised learning of functional network dynamics in resting state fMRI. Information Processing in Medical Imaging, Springer.

Eavani, H., T. D. Satterthwaite, R. E. Gur, R. C. Gur and C. Davatzikos (2013). Unsupervised learning of functional network dynamics in resting state fMRI. Information processing in medical imaging: proceedings of the... conference, NIH Public Access.

Eddy, S. R. (1996). "Hidden markov models." Current opinion in structural biology **6**(3): 361-365.

Finn, E. S., X. L. Shen, D. Scheinost, M. D. Rosenberg, J. Huang, M. M. Chun, X. Papademetris and R. T. Constable (2015). "Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity." Nature Neuroscience **18**(11): 1664-1671.

Fox, M. D., M. Corbetta, A. Z. Snyder, J. L. Vincent and M. E. Raichle (2006). "Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems." Proceedings of the National Academy of Sciences **103**(26): 10046-10051.

Fox, M. D., A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen and M. E. Raichle (2005). "The human brain is intrinsically organized into dynamic, anticorrelated functional networks." Proceedings of the National Academy of Sciences of the United States of America **102**(27): 9673-9678.

Fox, M. D., D. Zhang, A. Z. Snyder and M. E. Raichle (2009). "The global signal and observed anticorrelated resting state brain networks." J Neurophysiol **101**(6): 3270-3283.

Glasser, M. F., T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann and M. Jenkinson (2016). "A multi-modal parcellation of human cerebral cortex." Nature.

Glasser, M. F., S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster and J. R. Polimeni (2013). "The minimal preprocessing pipelines for the Human Connectome Project." Neuroimage **80**: 105-124.

Graves, A. (2012). Offline arabic handwriting recognition with multidimensional recurrent neural networks. Guide to OCR for Arabic scripts, Springer**:** 297-313.

Griffanti, L., G. Salimi-Khorshidi, C. F. Beckmann, E. J. Auerbach, G. Douaud, C. E. Sexton, E. Zsoldos, K. P. Ebmeier, N. Filippini and C. E. Mackay (2014). "ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging." Neuroimage **95**: 232-247.

Hanke, M., Y. O. Halchenko, P. B. Sederberg, S. J. Hanson, J. V. Haxby and S. Pollmann (2009). "PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data." Neuroinformatics **7**(1): 37-53.

Hutchison, R. M., T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D. A. Leopold, F. de Pasquale, O. Sporns, M. Walter and C. Chang (2013). "Dynamic functional connectivity: promise, issues, and interpretations." Neuroimage **80**: 360-378.

Ji, B., Z. Li, C. Coles, J. Kable, R. Zhang and X. Hu (2015). REDUCTION OF FUNCTIONAL CONNECTIVITY IN ADOLESCENTS PRENATALLY EXPOSED TO ALCOHOL. ALCOHOLISM-CLINICAL AND EXPERIMENTAL RESEARCH, WILEY-BLACKWELL 111 RIVER ST, HOBOKEN 07030-5774, NJ USA.

Jones, L. M., A. Fontanini, B. F. Sadacca, P. Miller and D. B. Katz (2007). "Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles." Proceedings of the National Academy of Sciences **104**(47): 18772-18777.

Juang, B.-H., W. Hou and C.-H. Lee (1997). "Minimum classification error rate methods for speech recognition." IEEE Transactions on Speech and Audio processing **5**(3): 257-265.

Keilholz, S. D. (2014). "The neural basis of time-varying resting-state functional connectivity." Brain connectivity **4**(10): 769-779.

Kingma, D. and J. Ba (2014). "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980.

Larkby, C. and N. Day (1997). "The effects of prenatal alcohol exposure." Alcohol Research and Health **21**(3): 192.

Lee, H.-L., B. Zahneisen, T. Hugger, P. LeVan and J. Hennig (2013). "Tracking dynamic resting-state networks at higher frequencies using MR-encephalography." Neuroimage **65**: 216-222.

Lee, M. H., C. D. Smyser and J. S. Shimony (2013). "Resting-state fMRI: a review of methods and clinical applications." American Journal of Neuroradiology **34**(10): 1866-1872.

Lindquist, M. A., C. Waugh and T. D. Wager (2007). "Modeling state-related fMRI activity using change-point theory." Neuroimage **35**(3): 1125-1141.

Liu, W., S. P. Awate, J. S. Anderson and P. T. Fletcher (2014). "A functional network estimation method of resting-state fMRI using a hierarchical Markov random field." NeuroImage **100**: 520-534.

Liu, X., C. Chang and J. H. Duyn (2013). "Decomposition of spontaneous brain activity into distinct fMRI co-activation patterns." Frontiers in Systems Neuroscience **7**: 101.

Liu, X. and J. H. Duyn (2013). "Time-varying functional network information extracted from brief instances of spontaneous brain activity." Proceedings of the National Academy of Sciences **110**(11): 4392-4397.

Lv, J., X. Jiang, X. Li, D. Zhu, H. Chen, T. Zhang, S. Zhang, X. Hu, J. Han and H. Huang (2013). Identifying functional networks via sparse coding of whole brain FMRI signals. Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on, IEEE.

Maaten, L. v. d. and G. Hinton (2008). "Visualizing data using t-SNE." Journal of Machine Learning Research **9**(Nov): 2579-2605.

Majeed, W., M. Magnuson, W. Hasenkamp, H. Schwarb, E. H. Schumacher, L. Barsalou and S. D. Keilholz (2011). "Spatiotemporal dynamics of low frequency BOLD fluctuations in rats and humans." Neuroimage **54**(2): 1140-1150.

Malisza, K. L., A.-A. Allman, D. Shiloff, L. Jakobson, S. Longstaffe and A. E. Chudley (2005). "Evaluation of spatial working memory function in children and adults with fetal alcohol spectrum disorders: a functional magnetic resonance imaging study." Pediatric research **58**(6): 1150-1157.

Mazziotta, J. C., A. W. Toga, A. Evans, P. Fox and J. Lancaster (1995). "A Probabilistic Atlas of the Human Brain - Theory and Rationale for Its Development." Neuroimage **2**(2): 89-101.

Mikolov, T., M. Karafiat, L. Burget, J. Cernocky and S. Khudanpur (2010). "Recurrent neural network based language model." 11th Annual Conference of the International Speech Communication Association 2010 (Interspeech 2010), Vols 1-2: 1045-1048.

Mitra, A., A. Z. Snyder, T. Blazey and M. E. Raichle (2015). "Lag threads organize the brain's intrinsic activity." Proc Natl Acad Sci U S A **112**(17): E2235-2244.

Mohr, P. N. and I. E. Nagel (2010). "Variability in brain activity as an individual difference measure in neuroscience?" J Neurosci **30**(23): 7755-7757.

Ou, J., L. Xie, C. Jin, X. Li, D. Zhu, R. Jiang, Y. Chen, J. Zhang, L. Li and T. Liu (2014). "Characterizing and Differentiating Brain State Dynamics via Hidden Markov Models." Brain topography: 1-14.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss and V. Dubourg (2011). "Scikit-learn: Machine learning in Python." The Journal of Machine Learning Research **12**: 2825-2830.

Power, J. D., A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Miezin and B. L. Schlaggar (2011). "Functional network organization of the human brain." Neuron **72**(4): 665-678.

Power, J. D., M. Plitt, T. O. Laumann and A. Martin (2017). "Sources and implications of whole-brain fMRI signals in humans." Neuroimage **146**: 609-625.

Rabiner, L. (1989). "A tutorial on hidden Markov models and selected applications in speech recognition." Proceedings of the IEEE **77**(2): 257-286.

Rabinovich, M. I. and M. Muezzinoglu (2010). "Nonlinear dynamics of the brain: emotion and cognition." Physics-Uspekhi **53**(4): 357.

Robinson, E. C., S. Jbabdi, M. F. Glasser, J. Andersson, G. C. Burgess, M. P. Harms, S. M. Smith, D. C. Van Essen and M. Jenkinson (2014). "MSM: a new flexible framework for multimodal surface matching." Neuroimage **100**: 414-426.

Robinson, L. F., T. D. Wager and M. A. Lindquist (2010). "Change point estimation in multi-subject fMRI studies." Neuroimage **49**(2): 1581-1592.

Sakoğlu, Ü., G. D. Pearlson, K. A. Kiehl, Y. M. Wang, A. M. Michael and V. D. Calhoun (2010). "A method for evaluating dynamic functional network connectivity and task-modulation: application to schizophrenia." Magnetic Resonance Materials in Physics, Biology and Medicine **23**(5-6): 351-366.

Salimi-Khorshidi, G., G. Douaud, C. F. Beckmann, M. F. Glasser, L. Griffanti and S. M. Smith (2014). "Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers." Neuroimage **90**: 449-468.

Setsompop, K., B. A. Gagoski, J. R. Polimeni, T. Witzel, V. J. Wedeen and L. L. Wald (2012). "Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty." Magn Reson Med **67**(5): 1210-1224.

Shakil, S., C.-H. Lee and S. D. Keilholz (2016). "Evaluation of sliding window correlation performance for characterizing dynamic functional connectivity and brain states." NeuroImage **133**: 111-128.

Shakil, S., C. H. Lee and S. D. Keilholz (2016). "Evaluation of sliding window correlation performance for characterizing dynamic functional connectivity and brain states." Neuroimage **133**: 111-128.

Siegel, J. S., A. Mitra, T. O. Laumann, B. A. Seitzman, M. Raichle, M. Corbetta and A. Z. Snyder (2016). "Data Quality Influences Observed Links Between Functional Connectivity and Behavior." Cereb Cortex.

Sikka, S., B. Cheung, R. Khanuja, S. Ghosh, C. Yan, Q. Li, J. Vogelstein, R. Burns, S. Colcombe and C. Craddock (2014). Towards Automated Analysis of Connectomes: The Configurable Pipeline for the Analysis of Connectomes (C-PAC). Front. Neuroinform. Conference Abstract: 5th INCF Congress of Neuroinformatics. doi: 10.3389/conf. fninf.

Smith, S. M., M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak and D. E. Flitney (2004). "Advances in functional and structural MR image analysis and implementation as FSL." Neuroimage **23**: S208-S219.

Smith, S. M., K. L. Miller, S. Moeller, J. Xu, E. J. Auerbach, M. W. Woolrich, C. F. Beckmann, M. Jenkinson, J. Andersson and M. F. Glasser (2012). "Temporally-independent functional modes of spontaneous brain activity." Proceedings of the National Academy of Sciences **109**(8): 3131-3136.

Spadoni, A. D., A. D. Bazinet, S. L. Fryer, S. F. Tapert, S. N. Mattson and E. P. Riley (2009). "BOLD response during spatial working memory in youth with heavy prenatal alcohol exposure." Alcoholism: Clinical and Experimental Research **33**(12): 2067-2076.

Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov (2014). "Dropout: A Simple Way to Prevent Neural Networks from Overfitting." Journal of Machine Learning Research **15**: 1929-1958.

Taghia, J., S. Ryali, T. Chen, K. Supekar, W. Cai and V. Menon (2017). "Bayesian switching factor analysis for estimating time-varying functional connectivity in fMRI." NeuroImage **155**: 271-290.

Talairach, J. and P. Tournoux (1988). Co-planar stereotaxic atlas of the human brain : 3-dimensional proportional system : an approach to cerebral imaging. Stuttgart ; New York, G. Thieme ; Thieme Medical Publishers.

Thompson, G. J., W. J. Pan, M. E. Magnuson, D. Jaeger and S. D. Keilholz (2014). "Quasi-periodic patterns (QPP): large-scale dynamics in resting state fMRI that correlate with local infraslow electrical activity." Neuroimage **84**: 1018-1031.

Trapp, C., K. Vakamudi and S. Posse (2017). "On the detection of high frequency correlations in resting state fMRI." Neuroimage.

Van Essen, D. C., M. F. Glasser, D. L. Dierker, J. Harwell and T. Coalson (2012). "Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases." Cerebral Cortex **22**(10): 2241-2262.

Van Essen, D. C., S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil and W.-M. H. Consortium (2013). "The WU-Minn human connectome project: an overview." Neuroimage **80**: 62-79.

Viterbi, A. J. (1967). "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm." Information Theory, IEEE Transactions on **13**(2): 260-269.

Wong, C. W., V. Olafsson, O. Tal and T. T. Liu (2013). "The amplitude of the resting-state fMRI global signal is related to EEG vigilance measures." Neuroimage **83**: 983-990.

Wozniak, J. R., B. A. Mueller, R. L. Muetzel, C. J. Bell, H. L. Hoecker, M. L. Nelson, P. N. Chang and K. O. Lim (2011). "Inter-Hemispheric Functional Connectivity Disruption in Children With Prenatal Alcohol Exposure." Alcoholism: Clinical and Experimental Research **35**(5): 849-861.

Yang, Z., S. LaConte, X. Weng and X. Hu (2008). "Ranking and averaging independent component analysis by reproducibility (RAICAR)." Human brain mapping **29**(6): 711-725.

Yeo, B. T., F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei and J. R. Polimeni (2011). "The organization of the human cerebral cortex estimated by intrinsic functional connectivity." Journal of neurophysiology **106**(3): 1125-1165.